DOCUMENT RESUME

ED 031 314 PS 002 103

By-Hillery, Milton C.; And Others

A Descriptive Study of Cognitive and Affective Trends Differentiating Selected Groups of Pre-School Children.

Wisconsin Univ., Milwaukee.

Spons Agency-Office of Economic Opportunity, Washington, D.C.

Report No-OEO-1401

Pub Date 69

Note-169p.

EDRS Price MF-$0.75 HC-$8.55

Descriptors-Cognitive Development, Cognitive Measurement, Compensatory Education Programs, Concept Formation, Culturally Advantaged, *Culturally Disadvantaged, Emotional Development, *Individual Characteristics, Language Development, Perceptual Development, *Preschool Children, *Program Evaluation, Rural Urban Differences, Self Concept, *Testing Programs

Identifiers-*Head Start

The purpose of the study was to describe ways in which disadvantaged children differ from their more advantaged peers in the areas of cognitive and affective developmental patterns, and to use this description to restructure curricular experiences for disadvantaged children. Five groups of approximately 30 preschool children each were tested three times with a battery of instruments. The instrument package, broken down into seven subsets, was designed to test a variety of developmental tasks. The appendixes describe the instruments used. Disadvantaged groups of children performed at lower levels in all measured areas of cognitive functioning, contrary to the investigators' expectations. Therefore, areas of serious deficit must be identified and tasks designed to build and improve these areas. The effects attributable to Head Start seem to be those changes which occur during the first weeks of any formal school program. Improvement in self-concept occurred in two of the five groups, which may be attributed to the integration of disadvantaged with advantaged children in those groups. Advantaged children are currently excluded from needed programs. Economic criteria provide an inaccurate basis for the allocation of funds to equalize educational opportunity, thus funding should be allocated according to the needs of diverse groups. (DR)

ED031314

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

F-OEO
OEO. 1401

A DESCRIPTIVE STUDY OF COGNITIVE AND AFFECTIVE TRENDS DIFFERENTIATING SELECTED GROUPS OF PRE-SCHOOL CHILDREN

Office of Economic Opportunity
Grant #1401

Milton C. Hillery
Ronald H. Lingren
Robert C. Remstad

University of Wisconsin at Milwaukee

1969

PS002103

The research reported herein was supported by the Research division of Project Head Start of the Office of Economic Opportunity, U.S. Department of Health, Education, and Welfare.

ACKNOWLEDGMENTS

The investigators wish to make several sincere acknowledgments to a number of individuals whose wholehearted cooperation made the project possible.

William H. Ashbaugh, Director of Department of Educational Research, G. Dwight Rowe, Research Coordinator, and Ruth Jensen, Elementary Education Supervisor, of the Milwaukee Public Schools and Superintendent Wayne Moldenhauer and Merton Peterson, Director of Elementary Education, of the Stevens Point Public Schools were the administrators who facilitated the project in their respective school districts. Our appreciation for their excellent cooperation and assistance is extended to the principals and kindergarten teachers in the seven elementary schools included in the project. The same note of appreciation is given to Miss Bonnie MacDonald of Wisconsin State University, Stevens Point, Director of the Head Start program in that community and her staff.

Finally, the authors' wish to express their gratitude to the many individuals who comprised the project staff. Their efforts were the actual source of this document.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS. ii

LIST OF TABLES iv

LIST OF FIGURES. viii

CHAPTER I: Nature of the Problem. 1

Introduction. 1
Objectives of the Study 4

CHAPTER II: Review of Related Literature. 6

CHAPTER III: Research Procedures and Methodology. 17

Introduction. 17
Selection of Groups to be Studied 18
Summary of the Instruments Used and the Administrative Schedule. 25
Data Collection Procedures. 29

CHAPTER IV: Results 33

Intergroup Comparisons. 34
Data Relevant to Individual Instruments 72
Summary 92

CHAPTER V: Interpretation of Results. 94

Summary 113

REFERENCES 117

APPENDICES 122

Appendix A, Description of Instruments. 123
Appendix B, Item Data for Wechsler Pre-School and Primary Scale of Intelligence. 154

LIST OF TABLES

Table | Page
1 Number of children being studied, arranged by sample groups, sex and age at date of fall testing. 24
2 Mean scores on the Columbia Mental Maturity Scale for five groups of children during four phases of testing 38
3 Mean scores on the Figure-Ground subtest of the Frostig DTVP for five groups of children during the four phases of testing 39
4 Mean scores on the Spatial Relations subtest of the Frostig DTVP for five groups of children during the four phases of testing 40
5 Mean scores on the Haptic Visual Matching Test for five groups of children during the four phases of testing. 41
6 Mean scores on the Picture Vocabulary subtest of the French PTI for five groups of children during four phases of testing 43
7 Mean scores on the Visual-Motor Association subtest of the ITPA for five groups of children during four phases of testing 44
8 Mean scores on the Visual Decoding subtest of the ITPA for five groups of children during four phases of testing 45
9 Mean scores on the Auditory Decoding subtest of the ITPA for five groups of children during four phases of testing 46
10 Mean scores on the Symbol Recognition Test, Part I, for five groups of children during four phases of testing. 47
11 Mean scores on the Symbol Recognition Test, Part II, for five groups of children during four phases of testing. 48

Table Page

12 Mean scores on the Size and Number subtest of the French PTI for five groups of children during four phases of testing. 51

13 Mean scores on the Elkind Measurement of Quantitative Comparisons for five groups of children given during the spring semester of kindergarten. 53

14 Proportions of subjects in five groups responding correctly to three types of extensive comparison items on the Elkind Measurement of Quantitative Comparisons instrument 54

15 Mean scores on the General Information Test for five groups of children during four phases of testing. 56

16 Mean raw scores and intelligence quotients on the Pintner-Cunningham Primary Test for five groups of children tested during the spring or fourth phase of testing 57

17 Mean scores on the Caldwell Preschool Inventory for five groups of children during the winter of kindergarten year. 58

18 Mean raw scores and standard deviations on ten subtests of the WPPSI for each of five subgroups . 60

19 Mean scale scores on ten subtests and the Verbal, Performance and Total scales of the WPPSI for each of five subgroups. 61

21 Mean response times (in seconds) on the Haptic Visual Matching Test for five groups of children during four phases of testing. 64

22 Mean frequencies of categorizations on basis of color, size or shape on the Block Sorting Task for five groups of children. 65

23 Mean scores on the Illinois Index of Self Derrogation for five groups of children during four phases of testing. 67

Table Page

24 Proportions of groups choosing the delayed reinforcement on the Delay of Gratification measure. For five groups of children during three phases of testing. 68

25 Mean frequencies of responses attributed to five members of family in five feeling categories on the Family Relations Test. Data for five groups from Phase I or II administration 70

26 Mean frequencies of responses attributed to five members of family in five feeling categories on the Family Relations Test. Data for five groups from Phase IV administration 71

27 Means and standard deviations of four tests in the Ability to Make Discriminations category for four phases of testing. 74

28 Mean scores on four tests in the category, Ability to Make Discriminations, for five age groups during three phases of testing. 75

29 Correlations between pairs of tests in the Ability to Make Discriminations category for Phases 2 and 4 76

30 Means and standard deviations of five tests in the Basic Comprehension Skills category for four phases of testing 77

31 Reliability estimates of five tests in the Basic Comprehension Skills category. 78

32 Correlations between pairs of tests in the Basic Comprehension Skills category for Phases 2 and 4 . 79

33 Mean scores on five tests in the category, Basic Comprehension Skills, for five age groups during three phases of testing. 80

34 Means and standard deviations on the French PTI, Size and Number subtest in the Basic Quantitative Skills category for four phases of testing 81

Table		Page
35	Mean scores on the French PTI, Size and Number subtest for five age groups during four phases of testing.	81
36	Means and standard deviations on the General Information Test for four phases of testing. . . .	83
37	Means of five age groups on the General Information Test for three phases of testing	83
38	Intercorrelations among the ten subtests of the WPPSI. (Based on 54 children in the urban, disadvantaged Negro groups.).	85
39	Mean scale scores on all subtests of WPPSI and mean Verbal, Performance and Full Scale IQ scores for the two disadvantaged groups	88
40	Mean response times on the Haptic Visual Matching Test and the Matching Familiar Figures Test in the Cognitive Style category during four phases of testing.	90
41	Means and standard deviations on two tasks in the Affective Traits category for four phases of testing.	91
42	Mean proportions or scores on two techniques in the Affective Trait category for five age groups .	91
43	Standardized mean scores for five groups of children on eleven instruments during final phase of testing.	97
44	Standardized mean scores for five groups of children on seven instruments during final phase of testing testing.	98

LIST OF FIGURES

Figure		Page
1	Graphs of mean scores on the Columbia Mental Maturity Scale for five groups at four testings	38
2	Graphs of the mean scores on the Figure-Ground subtest of the Frostig DTVP for five groups of children during the four phases of testing	39
3	Graphs of the mean scores on the Spatial Relations subtest of the Frostig DTVP for five groups at four testings.	40
4	Graphs of mean scores on the Haptic Visual Matching Test for five groups of children during the four phases of testing	41
5	Graphs of mean scores on the Picture Vocabulary subtest of the French PTI for five groups of children during four phases of testing	43
6	Graphs of mean scores on the Visual-Motor Association subtest of the ITPA for five groups of children during four phases of testing.	44
7	Graphs of mean scores on the Visual Decoding subtest of the ITPA for five groups of children during four phases of testing.	45
8	Graphs of mean scores on the Auditory Decoding subtest of the ITPA for five groups of children during four phases of testing.	46
9	Graphs of mean scores on the Symbol Recognition Test, Part I, for five groups of children during four phases of testing	47
10	Graphs of mean scores on the Symbol Recognition Test, Part II, for five groups of children during four phases of testing	48
11	Graphs of mean scores on the Size and Number subtest of the French PTI for five groups of children during four phases of testing	51

Figure Page

12 Graphs of proportions of children in five groups successfully completing three types of conservation tasks on the Elkind Measurement of Quantitative Comparisons 54

13 Graphs of mean scores on the General Information Test for five groups of children during four phases of testing 56

14 Graphs of score profiles of five groups on four subtests of the Caldwell Preschool Inventory. . . . 58

15 Graphs of mean scale scores on the five Verbal subtests of the WPPSI for five groups of children. (Administered during final month of kindergarten. . 62

16 Graphs of mean scale scores on the five Performance subtests of the WPPSI for five groups of children. (Administered during final month of kindergarten. . 62

17 Graphs of mean total scale scores on the Verbal and Performance Scales of the WPPSI for five groups of children. 63

18 Mean frequencies of color, size and shape categorizations on Block Sorting Task by five groups of children 65

19 Graphs of mean scores on the Illinois Index of Self Derrogation for five groups of children during four phases of testing 67

20 Graphs of proportions choosing the delayed reinforcement on the Delay of Gratification measure for five groups of children during three phases of testing . 68

21 Graphs of standardized mean scores of the five groups of children on four tests in the Ability to Make Discriminations category 99

22 Graphs of standardized mean scores of the five groups of children on the five tests in the Basic Comprehension Skills category. (During Phase 4 of testing 99

Figure Page

23 Graphs of standardized mean scores on proportions for the five groups of children on the two tests in the Quantitative Skills category. (During final phase of testing.). 100

24 Graphs of standardized mean scores of the five groups of children on the four tests in the General Ability and Knowledge category. 100

25 Graphs of standardized means and proportions for the five groups of children on the Affective Traits category tests. (During final phase of testing.) . 101

CHAPTER I

NATURE OF THE PROBLEM

Introduction

Within recent years the American people and their governmental representatives have directed increasing attention to the multitude of problems spawned by poverty and prejudice. It has become increasingly clear that remediation of these problems requires more than mere elimination of symptoms. Changes need to be wrought within the economically and socially disadvantaged subcultures which have been permitted to develop in the United States. These changes in people themselves define the problem as an educational one.

One significant approach to the problem of bringing about change through education within the various disadvantaged segments of our population has been that collection of pre-school programs known as Project Head Start. These programs are based on the assumption that some children, having spent their pre-school years in impoverished environments, begin school with deficiencies which severely limit their potential for profiting from an adequate school program. Head Start programs aim at an enrichment of this pre-school period, thus counteracting some of the deficiencies resulting from home and community background.

Unfortunately, Head Start programs have not had a solid foundation of knowledge about the development patterns of disadvantaged children and the ways in which they differ from those of the non-disadvantaged child upon which to build. While recognition has been made of the symptoms of the disadvantaged which set these youngsters apart from those of their middle class neighbors, relatively little is known of the differences in cognitive and affective development which may cause these symptoms. As a result, the curricular aspects of Head Start programs have necessarily been patterned after those of existing pre-school programs, either nursery schools or four year kindergartens. This was emphasized in the statement of aims and objectives in the pamphlet, *Program Head Start, Daily Program 1*:

> ". . .a Child Development Center should have the same aims for (the children's) pre-school experience in the Center as it would for children in any school anywhere."

This carries the implication that the deficiencies of the disadvantaged are quantitative rather than qualitative, i.e., that more experience, rather than a different kind of experience, is needed. However, it would seem that the very existence of Project Head Start is a negation of this point of view. That inability to profit from school experience can be remedied by more or earlier experiences of the same type is an argument of doubtful validity.

While it is true that increasing experience with the continued running of Head Start programs is beginning to generate for individual teachers or within single programs a repertoire of better, or more successful, techniques for working with disadvantaged children, this is a most inefficient way of building effective programs. What is needed for the efficient and effective building of curricular materials for Head Start and similar compensatory programs is adequate description of the ways in which disadvantaged children differ from their more advantaged peers. This study is an attempt to provide more adequate description of these differences.

The basic assumptions upon which this project was based are:

1. Many children from disadvantaged home environments develop cognitive and affective traits during the first five years of life which set them apart in many ways from non-disadvantaged children and which result in their inability to profit fully from school experiences designed to meet the needs of the non-disadvantaged.
2. Knowledge of the cognitive and affective characteristics of pre-school disadvantaged children is limited.
3. There is need for a descriptive study designed to identify specific cognitive and affective developmental traits of these children which can lead to the construction of curricular experiences which emphasize the strengths in these traits and attempt to remediate the weaknesses or deficiencies.
4. Any meaningful evaluation of Head Start projects must continue into the kindergarten year.

The goal of this study is the description of cognitive and affective developmental patterns of four and five year old children, based on the assumptions stated above. Specifically, both disadvantaged and non-disadvantaged children were observed and compared and the effects of Head Start programs upon these developmental patterns studied.

Objectives Of The Study

The goal of the description of cognitive and affective developmental patterns of four and five year old children is an ambitious one, one which is impossible of achieving in any complete sense. Practical restrictions on the success of meeting such a goal take the form of limitations on the amount of time which can reasonably be spent testing, measuring or observing individual children and on the size and representativeness of the samples of children to be studied. Within the range of these limitations, which have been recognized and stated in the proposal, the following objectives are outlined:

1. A description of both urban and small city-rural disadvantaged and non-disadvantaged children on a battery of tests, tasks and scales designed to measure a variety of cognitive and affective traits.
2. Estimation of within-group developmental trends over the eighteen to twenty-one month range of age usually found in Head Start and kindergarten children.
3. A description of the effect of the first year of school upon these traits through a series of repeated measurements.
4. Estimation of the effect of Project Head Start on these traits through the inclusion of a control group

of disadvantaged kindergarten children who did not participate in Head Start pre-school programs.

5. A comparison of the developmental patterns of disadvantaged and non-disadvantaged children in both urban and small town-rural settings.
6. The modification of existing instruments and construction of new ones for the measurement of the specified cognitive and affective traits.

The ultimate goal of such research as collected in this study is the utilization of the developmental knowledge gained to design curricular experiences, structured and timed in the most efficacious manner to compensate for the environmental experiential deficiencies of disadvantaged children and to capitalize on positive qualities disclosed by the developmental data. Thus, a final objective of this project was:

7. The development of a list of recommendations to be considered in the development of new curricular programs, both for pre-school and primary age disadvantaged children.

PS 002103

CHAPTER II

REVIEW OF RELATED LITERATURE

In theory, in research, and in practical application, the domain of the psychology concerned with concept formation has received significantly increased attention in the last two decades. Stemming from various psychological viewpoints, research has been directed toward the developmental patterns of the acquisition of concept formation skills; processes involved in concept attainment; factors external to the individual, or internal to the individual, affecting conceptualizing skills; and numerous other variations on the central theme. However, relatively few of these efforts have been directed towards studying the effects of environmental patterns and subsequent identification of basic conceptualizing skills or processes that are absent, or whose development has been delayed as a consequence of the environmental conditions.

A majority of efforts directed towards the evaluation of educational programs and the differential descriptions of population groups rely basically on general tests of intelligence or achievement.

The extent of socio-cultural bias in the mental testing instruments making up the bulk of school testing programs has been well documented over the last five to ten years. (Gordon, 1965 & 1963; Bean, 1942; Clarke and Clarke, 1953; Davidson, 1950; Higgins and Sivers, 1958; Siller, 1957; Karp and Sigel, 1965; Anastasi, 1961 & 1965; Duetsch, 1964 & 1965; Deutsch and Brown, 1964; Hunt, 1961;

Davis and Eells, 1953; Kidd, 1962; Whiteman, 1964; Dreger and Miller, 1960; Kimberg, 1963; Levinson, 1961 & 1963; McCandless, 1952; Pasamanick and Knoblock, 1955; Feldman and Weiner, 1964; Lehmann, 1959; Lennon, 1964; Lesser, Fifer and Clark, 1964; MacArthur and Elley, 1963; Mareton and Butcher, 1963; Noll, 1960; Riessman, 1962.)

However, such techniques provide ample measures of levels of attainment, amount of gain, and a quite reliable prediction of future successes in similar involvements. Such tests rely heavily upon concept formation skills, assessed relatively indirectly. Information gained from such techniques is largely of a global nature and at best answers the question of what the student is capable of doing with respect to certain identified areas of mental functioning. Little information can be gained as to determining why the individual was unable to perform successfully, or in a sense more relevant to this study, why the individual was unable to form the appropriate concepts in order to behave in the appropriate or required way. Piaget and Inhelder (1947) criticized metric scales such as those of intelligence and general achievement tests in a similar vein:

> "A test only gives us results on efficiency of mental activity without grasping the psychological operations in themselves. . .the test provides the sum of successes and failures, which is the actual result of past activities and attainments, but it leaves untouched the way in which these have been reached."

The literature dealing with a shift in emphasis from quantitative measurement of intelligence to qualitative assessment represents a very small but growing proportion of the mass of relevant psychological and educational literature. In the authors' opinions, it has not commanded the attention relative to its significance.

It is the rationale of this study that by qualitatively analyzing the acquisition of conceptualizing skills in terms of the developmental nature of the basic processes involved, as well as inferring differences derived from quantitative assessments, needed insights into the problem identified above can be gained.

The data collected can best be described in three categories. The first category of data is that which is primarily derived from instruments and forms frequently utilized by the Head Start program for evaluating purposes. The second category of data was obtained from those instruments or techniques structured to assess the developmental levels for those processes basic to the acquisition of concept formation skills and conceptualizing behavior itself. These data provide the major focus for this study. The third category of variables assessed were those which have been referred to as affective or affective-cognitive variables theoretically related to concept-formation skills.

The major source of data in the first category will be procedures that are categorized as general or global measures of intelligence or achievement.

The second category of data is composed primarily of those structured to assess conceptualizing behavior. This set of data was collected in a series of testings, spaced over the time periods covered by this study, to provide developmental data. Though not administered in a strict pre- and post-test manner, the sequential nature of the assessments make it appropriate for evaluative purposes as well.

This study derives its theoretical constructs from the approach offered by Dr. Irving Sigel of the Merrill Palmer Institute. It should also be noted that Sigel's approach draws heavily, though by no means exclusively, upon the developmental theory of Jean Piaget.

As a means of defining the notion on concepts, Sigel emphasized two dimensions. First is the function of concepts about which the author states:

> ". . .concepts serve as crucial links between the environment and the individual. . .when man employs concepts, he thinks in terms of symbols and classes or categories, he begins to reduce ambiguity and imprecision.
>
> Associated with most concepts or categories, moreover, is a set of appropriate behaviors be they thoughts, motoric acts, or fantasies. Man thereby has a repertory of behaviors with which to deal effectively with the environment." (Sigel, 1964)

Thus, in terms of function, there is, for convenience of description, a dichotomy. On the one hand there is a system of classes or categories that provide organization to the individual's perception of reality, and, on the other hand, schemas of behavior that are the consequence of said organization.

The second basic dimension of concepts is that they are acquired.

> "Concepts are acquired through a complex set of processes . . .such processes as discrimination, perception, transposition, and generalization - all facilitated by language." (Sigel, 1964)

A major focus of this research is the identification of developmental levels for those processes identified by Sigel, among others, as the processes which are basic to the acquisition of concept formation skills.

Included in this category are assessments of:

1. children's general conceptualizing or classifying behavior with emphasis upon discrimination skills.
2. strategies children utilize to make discriminations.
3. perceptual skills.
4. perceptual skills in relation to the development of linguistic skills.
5. language development and levels of communication.
6. labeling skills.
7. number concepts.

The assessment of a child's general conceptualizing or classifying behavior involving color, shape, size, use, number, etc., provides a source of data for inferring discrimination skills. Discrimination primarily involves, in Sigel's description, the facts of learning that objects differ from one another and that objects have multiple characteristics and attributes. Gagne (1965), using the term "multiple discrimination", has described it as becoming capable of making different responses to the different members of a particular collection. With reference to discriminating behavior, Deutsch (1963) has found a pervasive inferiority among disadvantaged children. In terms of general conceptual thinking, Siller (1957) has found differences between socio-economic classes. Riesmann (1963) and Gordon (1964) have described limitations of the disadvantaged child's conceptualization. However, all of these sources refer primarily to older age groups than is the concern of the study.

An additional outcome of this study should be some knowledge of what strategies children utilize to make these discriminations and thus what strategy they use to categorize things. The work of Kagen, Moss and Sigel (1963) has led to the inference that there are three general strategies involved: 1) Descriptive or look alike; 2) Relational or Thematic - one object is commonly used in connection with another; and 3) Categorical or Symbolic - because of certain concrete or abstract attributes, objects belong together in a certain class or category. It is further inferred that such strategies emerge developmentally in the same order as they are listed. These authors have suggested several techniques for identifying which strategy or strategies predominate a child's conceptualizing processes.

Because the age of the children studied precludes the more advanced forms of conceptualizing, a major emphasis was placed on perception which is the major determinant of concepts at earlier developmental levels. A study by Boger (1952) has indicated the relevancy of perceptual skills, and the beneficial effect of training on these skills, to performance on general I.Q. tests. Similar results on a general Readiness Test were achieved as a part of an experimental program reported by Brazziel and Terrell (1962). Covington (1962) has indicated that upper-status children (status defined by level of education of parents) were superior to lower-status children in visual perceptual ability, but that the differences diminished with stimulus familiarization.

The study also dealt directly with perceptual processes as they are related to the development of linguistic skills, another process identified by Sigel as basic to the acquisition of conceptualizing skills. The assumption that perceptual processes and sensory modalities contribute differentially to conceptualizing skills and learning in general has been supported by the work of Irwin (1948), Deutsch and Brown (1964), Sievers (1965), C.P. Deutsch (1964) and Newton (1964).

To make the analyses of the above variables meaningful, it is also desirable to study language development directly, through some form of picture vocabulary technique.

There has been considerable research that would indicate that socio-cultural deprivation significantly affects language development in such a way as would be depicted by these "levels of communication". Berstien (1961) suggests that lower-class children do not learn the "formal" linguistic code which allows for elaboration of meaning and suggests the possibilities coherent in a complex conceptual hierarchy for the organizing of experience. Thus, this deprivation impedes or restricts cognitive development. Though the results at early school age were not conclusive, John (1963) concluded that "acquisition of more abstract and integrative language seems to be hampered by the living conditions in the homes of lower-class children". Opportunities for learning to integrate are rare.

Deutsch's work (1964) has likewise pointed to developmental disparities between social classes in language development, with the differences being most pronounced on measures reflecting abstract and categorical use of language.

Another area of assessment directly related to concept formation is that of number concepts. Not only do these concepts serve as an example of behavioral schema emanating from acquired concepts, but are also a form of mental functioning represented in a substantial proportion of educational objectives.

The third, and final category of variables assessed are those which have been referred to as affective, or affective-cognitive variables theoretically related to concept-formation skills. In general each of the variables discussed are related in the sense that they effect the development of conceptualizing skills and/or the manner in which one is able to utilize the concepts that have been attained.

The affective variables would include some measure of self-perception. Evidence for the relationship between self-perception and school achievement and learning in general, which would include acquisition of concept and conceptualizing skill, have been reported in Ausubel, D.P. and Ausubel, Pearl (1963). Research reports as summarized by Dreger and Miller (1960) would indicate that self-concepts are markedly less adequate or less positive for Negroes than for whites. The relationship of self-concept to academic achievement for the disadvantaged has also been investigated by Edwards and Webster (1963).

It is also proposed to assess the individual's capacity for delay of gratification. Evidence supporting the hypothesis that children from lower socio-economic levels are characterized by a low capacity for delay of gratification could be inferred from the findings of Le Shan (1952) who maintained that lower-lower class families tend to train children with immediate punishments and rewards. Families of the middle and lower-upper classes tend to stress the future, and punishment and reward are often deferred. Bandura and Walters (1963) have pointed out that: "The transmission of self-indulgent patterns may be associated with a low level of technology and precarious economic and social life, which persist in spite of contact with a more provident social life."

The identification and description of significant characteristics of the children's family and peer relationships is also investigated. References, descriptions and research results are numerous for this aspect of the topic of disadvantaged children. (Brofenbrenner, 1961; Keller, 1963; Casler, 1961; Riessman, 1963; Cloward and Jones, 1963; Deutsch, M., 1964; Douglas, 1964; Maas, 1951; Dreger and Miller, 1960; Kohn, 1959; Goff, 1954; Neugarten, 1946; Stendler, 1951.) Unfortunately relatively few of these efforts have been integrated with qualitative analyses with respect to the development of conceptualizing skills. Gordon (1963, 1965) has stressed the need for such appraisals.

The work of Kagen, et.al. (1964) and Kagen (1965) on information processing in the child is relevant to the purposes of this study. The behavioral phenomena examined in these studies would seem to be in that overlapping area of affective and cognitive behavior, frequently classified as cognitive styles. This work has led to the inference that individual variance in cognitive behavior can in part be accounted for by differences in informational processing behavior. The work relevant to this study has focused on the "reflectivity-impulsivity" dimension which Kagen identifies as being primarily a developmental characteristic. Thus the relationship between an individual's relative position on the reflectivity-impulsivity dimension and his developmental level of conceptualizing skills and facility with attained concepts is a potential source of important insight, as is his general performance on the pre-school achievement battery. Another question examined is: "Is position on the continuum an attribute which relatively consistently discriminates the disadvantaged from the non-disadvantaged?"

The literature cited above provided a theoretical base of intellectual development and further identified the component parts of that development. The existing literature was then reviewed to identify that sub-set of these components or aspects of intellectual development for which evidence existed, either empirical or logical, that disadvantaged children had developmental deficits. The literature revealed that although some such studies had been done, they were fragmentary in terms of the total theoretical constructs. These findings reinforced the authors' contention that a more comprehensive investigation of specific developmental trends should be undertaken to further qualitatively describe differences, or lack thereof, between the disadvantaged and their peers, and to provide a means for evaluating the nature of the interrelationship among the developmental factors identified from the literature.

CHAPTER III

RESEARCH PROCEDURES AND METHODOLOGY

Introduction

The general strategy followed in the execution of this study was:

1. A package of instruments for measuring a variety of developmental traits was devised requiring approximately two and a half hours of individual testing time.
2. Five groups of approximately 30 children each were selected for study representing urban-rural disadvantaged, disadvantaged-non-disadvantaged, and Head Start-non-Head Start dichotomies.
3. Each youngster was measured on the complete battery of instruments three times (four times for those groups in summer Head Start programs) during the period from June 1966 to June 1967. Measurements were made of each child within two weeks of the day on which his chronological age was some multiple of three months. This resulted in measurements at constant chronological ages while allowing school experience to vary within a range of plus or minus six weeks.
4. The procedures for analyzing and reporting of the

data were to be organized in such a way that comparisons could be made among the five groups of children, representing the above listed dichotomies, across the four phases of testing and among five different age groups.

Selection Of Groups To Be Studied

The description of categories of individuals based on repeated measurements over a period of time as long as a school year dictates that these individuals must belong to some intact groups rather than be a representative sample drawn together solely for the purpose of data collection. This is even more essential when most of the measurements are to be made by individually administered psychological instruments, and the children are so young as to require that the data collector go to the children rather than bringing the children to the data collector. Thus, the desire to study a truly representative (in the sense of a random or stratified random sample) sample of children representing the population of interest was irreconcilable with the practical requirement that the samples of individuals being studied be kept together for an entire school year.

The initial selection factor was that the sample groups be drawn from communities in which Head Start programs were conducted during the summer of 1966. Secondly, the groups were composed of children who would be beginning kindergarten in September of 1966. The third selection factor was the mechanical or logistic concerns described above which necessitated the identification of more or less intact groups which could be identified as the following:*

A. An urban "core area" group who had participated in a summer Head Start program.

B. An urban "core area" group who had not participated in a summer Head Start program.

C. An urban "non-core area" group not eligible for Head Start.

D. A "small city-rural" group who had participated in a summer Head Start program.

E. A "small city-rural" group of "non-disadvantaged" children.

The first three groups, identified with the cooperation of the Milwaukee, Wisconsin, Public Schools Division of Research, consisted entirely of black children. The remaining two groups, identified with the assistance of the Stevens Point, Wisconsin, Public Schools and Wisconsin State University-Stevens Point, Department of Home Economics, which operated the summer Head Start program in Stevens Point, consisted exclusively of white children. It is recognized that this resulted in a complete confounding of race with the urban-small city-rural dimension. For those who might be critical of this confounding, it should be pointed out that the study is a description of various groups of kindergarten age children and that the confounding experienced in this study is very much a reality in many areas of the northern United States.

---

*In the discussions which follow, the five groups will be identified as above by the letters A through E.

Actually, the selection process took place in two phases:

1. Groups A and D were selected in June from the enrollment lists of the Head Start programs in Milwaukee and Stevens Point.
2. The remaining groups were selected late in August and early in September from the kindergarten registration lists in the two cities.

The selection process differed somewhat between Milwaukee and Stevens Point for both phases. In Milwaukee it was necessary to first identify a school housing a summer Head Start program that met two criteria:

1. A neighborhood environment for the school which would reasonably lead to the assumption that all children attending the school during the regular school year could be considered "disadvantaged" children.
2. Sufficient numbers of children in attendance at the given Head Start center who would be entering kindergarten in the fall of the year to provide the desired sample of thirty.

Thus, a public school housing a Head Start program with four Head Start classes identified from which it was possible to select thirty children who would be entering kindergarten in the fall, either in that school or a neighboring elementary school that drew students from a similar impoverished neighborhood. Due to the fact that the Milwaukee Public schools have a mid-semester matriculation policy, many of the children attending Head Start had either one semester of kindergarten or would not enter kindergarten until the second semester of the school year. This factor and the unexpected amount of transiency among this urban disadvantaged group made it impossible to attempt to balance the sample by age or sex exactly as had been proposed. However, as Table 1 (p.24) shows the resulting distributions were adequate with regard to sex. This matriculation policy also tended to restrict the age range of the Milwaukee sample groups in comparison to the Stevens Point groups. The mean age at entrance into kindergarten of the Milwaukee children tends to be younger than the Stevens Point group, 5 years, 4 months as compared to 5 years, 7 months.

Due to logistical problems and the desirability of keeping the project activities restricted to as few classrooms as possible, the later selection of the remaining two urban groups, Groups B and C, were restricted as was the first.

The urban disadvantaged group selected in the fall, Group B, those who had not taken summer Head Start or any other Head Start program, were selected from among the populations of the two schools selected as project schools as described above. Because only half of the kindergarten population fit the criteria of "starting kindergarten in the fall of 1966" there was, as indicated, little range of selection and the sample included most of the eligible population.

The choice of sample Group C, the urban non-disadvantaged group, was a relatively intact group. As has been indicated above, the racial composition of Head Start classes in Milwaukee and Stevens Point resulted in confounding the urban-small city dichotomy with race. To avoid further complications the decision was made that Group C should also be all black. An elementary school on the fringe of the "inner city" was selected. After eliminating children who had had Head Start or met the economic criteria, the sample chosen made the bulk of the eligible children. As will be discussed later, some questions must be raised concerning the adequacy of this group in meeting the original design of the study.

In selecting the initial summer group from the Stevens Point Head Start program several factors had to be considered. First, it was necessary to identify the schools which the children would be attending in the fall and identify those with the highest number of students. In this way it was hoped to restrict the project activity to as few schools as possible and yet have an adequate population from which to select the sample so as to obtain balance in age and sex. Since a large majority of the Head Start children were from rural areas they would be bussed to their kindergarten classes as they were to the school which housed the summer Head Start program. In most schools kindergarten children who were bussed came either mornings or afternoons. Due to the fact that the testing and related work was to be done by a single full time observer, it was necessary to balance morning and afternoon attendance. As a result of these factors, the sample was taken from children who would be attending three schools in the fall. One of these schools was scheduled to have Head Start participants in both morning and afternoon sessions. One of the other two would have Head Start children in the morning; the remaining school would have Head Start children in the afternoon.

The sample selected in the fall, Group E, the non-disadvantaged group, was selected from among the kindergarten classmates of the D group, with the contingency that there was to be no question as to their being non-disadvantaged. At no time was an attempt made to control for ruralness in either sample. Both included children who lived within the city-limits as well as outside. Because the kindergarten scheduling was primarily determined by bus schedules and related factors, the groups contained approximately equal numbers of rural and town children.

Because of the heterogenous nature of the classes in these schools, no assumptions about a child being disadvantaged or its antithesis could be made on the knowledge of the school's neighborhood as was the case in Milwaukee. The samples were selected and the cooperation of the county welfare office and county nurse's office were used to substantiate the appropriateness of the child's economic status.

Also, unlike the situation in Milwaukee, there were ample numbers of children from which to select the project samples. Thus, selection of children aimed at meeting three criteria: (a) balance of the groups as to sex, (b) balance the number within each group by age so that there were approximately an equal number of subjects with birthdates in each quarter of the year, and (c) spread the birthdates of the subjects across the days of the quarters so as to facilitate the testing schedule. Although it was not possible to satisfy all three criteria completely, the selected sample reasonably approximated the ideal.

Table 1 below shows the actual size and distribution by age and sex of the five groups at the start of school in September.

Table 1. Number of Children Being Studied, Arranged by Sample Groups, Sex and Age at Date of Fall Testing.

Group	Age At Fall Testing										Totals		
	5-0		5-3		5-6		5-9		6-0				
	B	G	B	G	B	G	B	G	B	G	B	G	Group
A	3	5	5	5	7	5	1	1	0	0	16	16	32
B	1	3	8	6	7	5	1	0	0	0	17	14	31
C	7	1	3	8	3	4	2	1	0	1	15	15	30
D	0	0	3	4	2	6	3	2	4	2	12	14	26
E	0	0	0	7	6	3	4	3	5	2	15	15	30
Totals													
Sex	11	9	19	30	25	23	11	7	9	5	75	74	
Age	20		49		48		18		14				149

With one exception, all the instruments or procedures used in this study were individually administered to the children according to the schedule described above. The one exception was the Pintner-Cunningham I.Q. Test. This had not originally been included in the battery of assessment procedures, but as it was administered to the

Milwaukee kindergarten children as a part of the school system's regular testing program, it was decided to administer it to the Stevens Point samples as well.

Summary Of The Instruments Used And The Administrative Schedule

The subjects of the study were to be tested on the battery of instruments three or four times during the course of the study. These times were during the weeks that the summer Head Start programs were being run and once each in the fall, winter and spring months of the school year. These four periods will be referred to as Phases in the following discussion, Phase I being the summer and Phase II, III and IV the three school year periods.

The original and ideal procedure would have been to administer all assessment techniques to all the subjects in all four or three phases in which they were to be tested. However, this was not the case for two major reasons. First, the timing of the grant award was such that the summer Head Start program was about to begin. Fortunately, the investigators had been involved in testing related to a separate evaluation program and the instruments utilized in this other study were a portion of the battery developed in preparing the proposal for this study. This enabled a fairly quick start once the personnel had been hired. However, the problems of getting the project running effectively and completing arrangements for the selection of the remaining three groups impeded the development and preparation of materials for additional procedures. Some of these were prepared during the summer and other during the fall. As a consequence some children in Groups A and D were not tested on all procedures during the summer phase. Several of the instruments prepared or developed during that time were administered to only a portion of the two sample groups.

All children in all five groups received the same basic battery during the fall or Phase II testing. Four additional procedures to be included in the battery were prepared or developed during the fall. At the same time it became apparent that the basic battery was not only taking as much time for each child as the investigators and school staff felt was desirable, but that it was also pushing the examiners' abilities to complete the testing within the required time, both for individuals and for the group as a total. Since it was felt that the examiners should have some time to work with children aside from the testing situation, both to build rapport with children and maintain good relationships with the teachers, it was decided that additional instruments could be administered only at the expense of some reduction in the basic package.

Given these conditions, it was decided to include the additional assessment procedures by a schedule which would hold the testing time per individual approximately the same as it had been but reduce the number in each sample group that would be administered each test during the remaining phases. It was felt that the information to be gained by the use of the additional instruments would be worth the loss in precision due to reduced sample sizes on the basic set of instruments. Thus, an additional number of instruments were only administered during the last two phases.

The following is a list of the various assessment procedures employed in the study. While these will be described in detail in a later section of this report, the list serves to explain the schedule used during the four phases of the study.

During the summer of 1966, the following assessment procedures were administered to all of the subjects in sample Groups A and D participating in summer Head Start programs in Milwaukee and Stevens Point.

1. General Information Test
2. Marianne Frostig Developmental Test of Visual Perception
    a. Figure-Ground subtest
    b. Spatial-Relations subtest
3. Columbia Mental Maturity Scale
4. Symbol Regognition Test
5. French Pictorial Test of Intelligence
    a. Picture Vocabulary subtest
    b. Size and Number subtest
6. Delay of Gratification
7. Illinois Test of Psycholinguistic Abilities
    a. Visual Decoding subtest
    b. Auditory Decoding subtest
    c. Visual Motor subtest
8. Bene-Anthony Family Relations Test

In addition to these techniques three other instruments were administered to portions of the children in Groups A and D during Phase I. These were:

9. Illinois Index of Self-Derrogation

   (Approximately half of Group A and all of Group D.)

10. Haptic

    (Over half of Group D.)

11. Matching Familiar Figures

    (Over half of Group D.)

During Phase II all of the above eleven instruments were administered to all children in the five groups with scattered exceptions due to prolonged absences or to refusal of children to cooperate. During this phase procedures for use of the three additional instruments listed below were developed and the examiners trained to administered them. These were:

12. Sigel Categorizing Task
13. Block Sorting Task
14. Elkind Measurement of Quantitative Comparisons

As mentioned earlier, the addition of these three tests to the testing package resulted in a total time which both teachers and investigators judged excessive. This problem was handled by subdividing the instruments on the basis of testing time into four subpackages of tests. During each of Phase III and Phase IV each child was administered three of these four subpackages according to a schedule which was designed to maintain a balance among the groups as to age.

Because of the importance that had been attached to the Caldwell Pre-School Achievement Test during the evaluation of the first summer of Head Start it was determined that it should be administered once during the study. As administration time for the Caldwell was approximately the same as for the Bene-Anthony Family Relation Test and as the Bene-Anthony was measuring attitudes toward family rather than a developmental trait, the Caldwell was administered during Phase III in place of the Bene-Anthony.

Finally, the Pintner-Cunningham Group Intelligence Test and the Wechsler Pre-School and Primary Scale of Intelligence (WPPSI) were administered to all children during Phase IV. In Milwaukee the Pintner-Cunningham was administered by the kindergarten teachers as a part of the citywide testing program. In Stevens Point it was administered by one of the principal investigators. The WPPSI was administered individually by a group of qualified school psychologists during the last weeks of school.

Data Collection Procedures

The logistics of collecting a package of measurements of the size described in the prior section on approximately 150 kindergarten children at three month intervals presents considerable problems. The investigators were particularly cognizant of the possible very disruptive influence the study could have had on the kindergarten classes being studied. The strategy followed was to make the data collection an integral part of the classroom operation. Rather than have a number of individuals attempting to test all the children being studied within a given short time span, a single individual made all of the measurements on all of the approximately thirty children in a particular group. This individual was placed in the Head Start or kindergarten class as a teacher assistant. While this assistant's primary responsibility was that of tester or observer for this study, a variable percentage of her time was at the disposal of the teacher for the facilitation of instruction. The testing program was a continuous one with each child being repeatedly tested at those chronological ages which are multiples of three months. Thus, during the first three months of kindergarten, every child went through the battery of measurements, being tested within a week of the day on which he was 60, 63, 66, 69 or 72 months old. While there were some difficulties in maintaining this schedule during the Head Start sessions due both to the lack of adequate preparation time and the short length of the session, it was closely followed during Phases II, III and IV. The advantages of such a procedure were believed to be:

1. Children had the opportunity to view the tester as a familiar, integral part of the class, thus reducing the intrusive effect of testing.
2. The tester's job was a continuing one over the duration of the study, thus facilitating the hiring of competent persons.
3. Being a permanent part of the class situation improved

the efficiency of the tester by eliminating the need for warm-up periods with individual children and providing flexibility of scheduling.

4. The tester was able to make some minimal contributions to the learning environment of the classrooms.

While it was recognized that having one person do all the testing for a given group might introduce a bias not present when a team of testers test all groups, the above advantages pointed to the procedures used.

The personnel selected for the positions of testers or observers were senior undergraduate students in education. The three half-time observers from Milwaukee were selected on the basis of their ability as demonstrated during an evaluation program of Milwaukee County Head Start programs early in the summer of 1966. The observer in Stevens Point was selected following interviews with a number of individuals qualified and available for the position. Because the disadvantaged and non-disadvantaged children were attending the same schools and were approximately evenly divided between morning and afternoon kindergarten it was advantageous to employ one observer full time for these two groups rather than two half-time employees as originally proposed.

All four of the observers received a thorough orientation in the instruments and techniques employed in the study and had practice sessions on non-sample children. The testers were regularly supervised by one of the principal investigators to maintain standardization of administration among the five groups.

The research design and procedures described in this chapter represent the attempt by the investigators to select a sample representative of the populations identified for investigation in the proposal, to arrange a testing schedule that provided for repeated individual administration of nearly all assessment procedures on a periodic basis keeping age of subjects controlled, and to provide for competent examiners who would have ample opportunity to establish the necessary rapport with the young children.

CHAPTER IV

RESULTS

The data to be presented in this chapter will be organized in several diverse formats to accomplish the purposes of the study. The primary purpose of the study was to identify, if possible, relevant dimensions of cognitive and affective development in Head Start and post-Head Start age children that would be of value in restructuring curricular experiences for disadvantaged children. To provide groups for comparison, five distinct groups of children were studied representing various combinations of the Head Start-non-Head Start, disadvantaged-non-disadvantaged, urban-small town and Negro-white dichotomies. As the intent is to describe each group on the entire battery of instruments as well as to point out relevant differences, the procedure of looking at the five groups simultaneously but in terms of smaller subsets of the instrument package will be followed.

A second purpose of this study was to provide researchers and curriculum builders with longitudinal developmental data on a variety of instruments designed to discriminate among children of varying cognitive skills and affective traits. For this purpose, data pertaining to individual tests and subtests of tests will be systematically presented, test by test, for the entire group of children studied. Gains with age and over time, correlations with other instruments, and reliability data will be presented.

Intergroup Comparisons

For the purposes of comparing the performances over time of the five groups studied, the package of instruments administered have been broken down into seven subsets measuring:

1. Ability to make discriminations.
2. Basic comprehension skills.
3. Basic quantitative skills.
4. General ability and knowledge.
5. Cognitive style.
6. Affective traits.
7. Family relationships.

While recognizing that these categories are not independent and that any given instrument taps dimensions of several of the categories, the instruments have been placed in categories which reflect the primary dimension of the task.

Results will be presented in the form of mean scores for each of the five groups during each of the four phases. Presentation will take two forms, tabular and graphical. For ease of interpretation each table and its accompanying graph is presented on a single page. For each instrument, a standard error of the mean has been calculated by dividing a pooled estimate of the standard deviation by the square root of the harmonic mean of the sample sizes on which the 17 group-by-phase means were calculated. While there are obvious problems in applying statistical tests to groups chosen as these were, there is a need to indicate in some fashion the precision of the estimates and to provide some yardstick against which the meaningfulness of differences between means can be measured. We have chosen the convention that differences as large as two standard errors are differences worthy of discussion. Each table contains the estimated standard error and the harmonic mean of the sample sizes upon which it was based. The corresponding table has a line, two standard errors long, drawn in the upper left-hand area of the figure for ease in interpreting group improvements over time as well as differences between groups.

Careful perusal of the table and corresponding graphical presentations will emphasize the regularities in trend lines that were obtained and the fairly stable differences between groups over time. Thus, it is not our intention to discuss at length all differences of two standard errors but instead will restrict our discussion to general statements of the regularities in the data and the occasional departures from regularity.

Category 1: Ability To Make Discriminations

Four tests or subtests of the instrument package were designed to measure ability of children to make a variety of visual discriminations. The Figure-Ground subtest of the Frostig test is designed to measure the child's ability to make perceptual shifts in identifying figures against increasingly complex grounds. The Spatial Relations subtest of the Frostig involves the making of simple form and shape discriminations. A more comprehensive instrument which introduces the need to classify as well as discriminate is the

Columbia Mental Maturity Scale. This test utilizes perceptual discriminations involving color, shape, size, use, number, kind and missing parts. The Haptic Visual Matching Test requires subjects to make three dimensional form discriminations through tactile sensory data.

Results of testing the five groups over the four phases of the study are reported in Tables 2, 3, 4 and 5 and in Figures 1, 2, 3 and 4 which appear on the following pages.

A pattern is established on these four instruments which tends to be repeated in all areas of testing. Although Groups A, B and D have in common the economic criteria for disadvantaged which made them eligible for Head Start programs, the clustering of groups in our data is not along these lines. The tendency, which shall be noted throughout the study, is for Groups A and B to cluster at a generally low level of performance, for Groups D and E to cluster at a measurably higher level of performance and for Group C to be a type of "swing" group, on some measures resembling the A, B cluster, on other measures the D, E cluster and on yet other measures being clearly in between. This differentiation into essentially three groups on the basis of task performance is an urban-small town type of split or perhaps a Negro-white split since this classification is completely confounded with the other dichotomy.

Examination of Table 2 and Figure 1 which presents the results of the Columbia Mental Maturity Scale reveals two distinct clusters with Groups C, D and E consistently performing five or more standard errors above Groups A and B. The gains over the six or nine month (depending upon the Group) period between first and last test administration are 16.6 and 14.7 for Head Start Groups A and D and 7.6, 8.7 and 7.8 for Groups B, C and E. Comparing these changes in raw score to corresponding mental ages results in average growth of approximately 9 and 11 months for Groups A and D and of 4, 8 and 7 months for Groups B, C and E.

The results of the Figure-Ground subtest of The Frostig DTVP (Table 3 and Figure 2) provides a pattern similar to the Columbia with Groups C, D and E clustered at the top and Groups A and B measurable lower. While the curve for Group A is about half-way between that for B and those for the other three, note that the mean for B at any given phase is approximately equal to that for A during the preceding phase. This is the result one would expect if the differences were school related and the summer Head Start program did indeed provide a head start which did not dissipate over the kindergarten year.

Table 4 and Figure 3 summarize the Spatial Relations subtest of the Frostig DTVP. The urban groups, A, B and C are clustered at the lower level with the means for phases II and IV being very close to or within the two standard error spread. Groups D and E are measurably higher. Note that the large increase for the three groups (C, D and E) attending non-core area schools during the first semester while the large increase for the groups in core area schools occurs during the second semester.

Table 2. Mean scores on the Columbia Mental Maturity Scale for five groups of children during four phases of testing. (Standard error = 2.21; $\bar{n}_h$ = 26.14)

Phase \ Group	A	B	C	D	E
I	24.2	-	-	37.7	-
II	29.6	26.5	43.3	46.4	48.6
III	33.2	32.5	49.9	51.8	50.7
IV	40.8	34.1	52.0	52.4	56.6

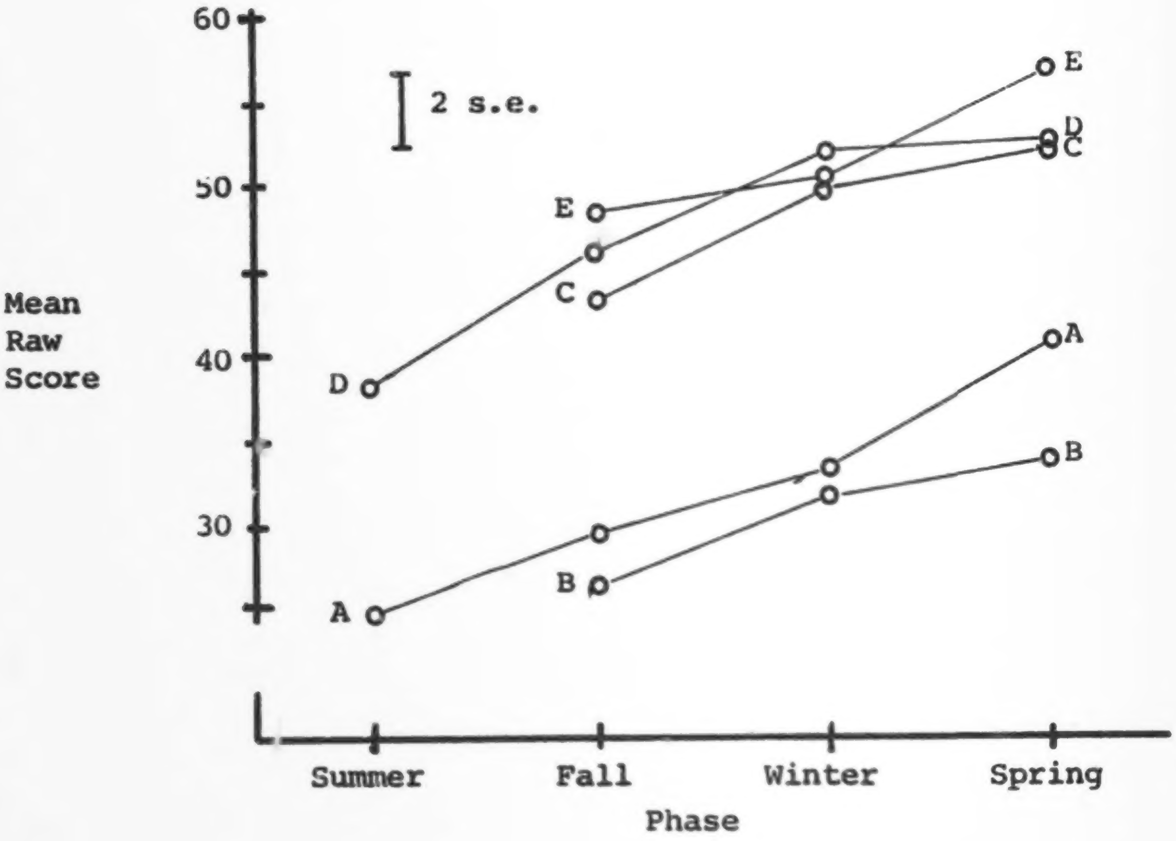

Figure 1. Graphs of mean scores on the Columbia Mental Maturity Scale for five groups at four testings.

Table 3. Mean scores* on the Figure-Ground subtest of the Frostig DTVP for five groups of children during the four phases of testing. (Standard error = 2.39; $\bar{n}_h$ = 22.07)

Phase \ Group	A	B	C	D	E
I	34.8	-	-	45.5	-
II	45.0	35.5	50.8	52.0	53.4
III	49.9	40.3	59.6	59.3	59.9
IV	54.2	49.2	64.9	59.4	60.0

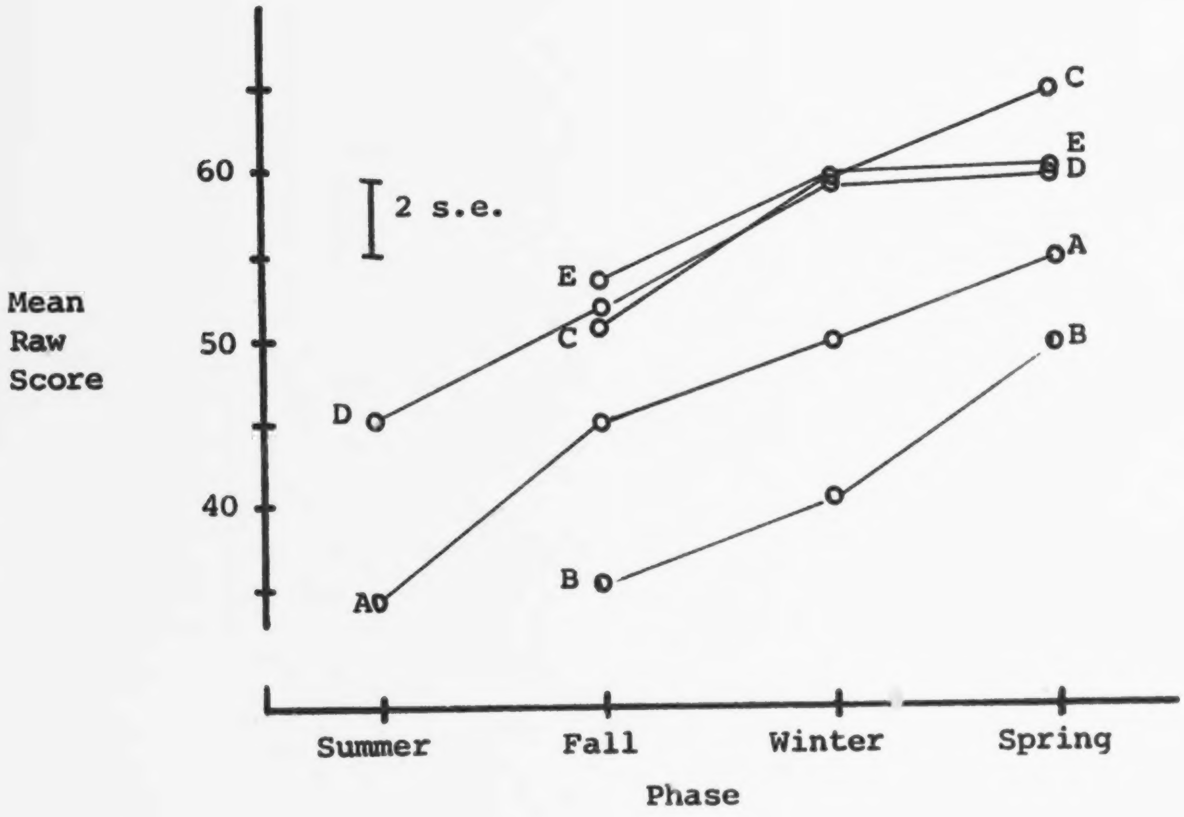

Figure 2. Graphs of the mean scores* on the Figure-Ground subtest of the Frostig DTVP for five groups of children during the four phases of testing.

*Based on scoring system developed for this study and described in Appendix.

Table 4. Mean scores* on the Spatial Relations subtest of the Frostig DTVP for five groups of children during the four phases of testing. (Standard error = 2.80; $\bar{n}_h = 22.07$)

Group Phase	A	B	C	D	E
I	20.0	-	-	29.0	-
II	25.4	20.4	26.1	33.1	38.3
III	24.3	19.8	33.8	44.0	50.0
IV	29.7	32.0	33.8	43.6	50.7

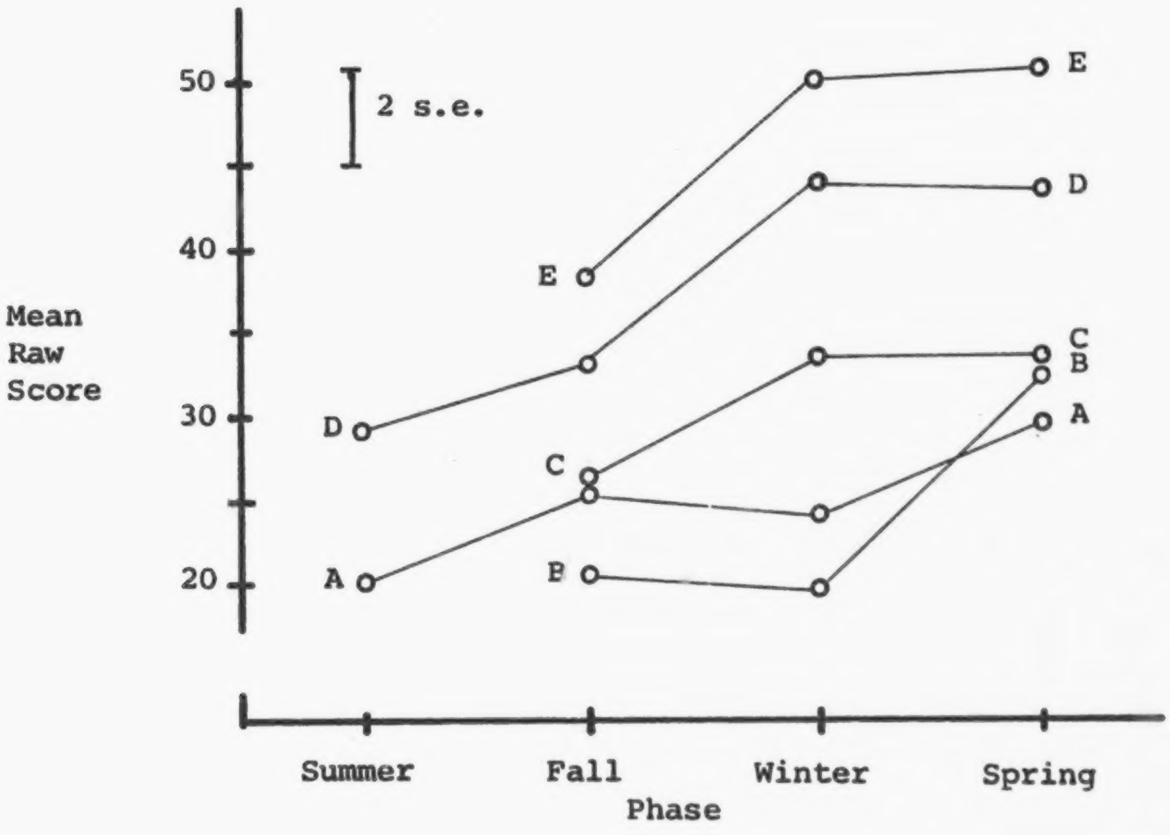

Figure 3. Graphs of the mean scores* on the Spatial Relations subtest of the Frostig DTVP for five groups at four testings.

---

*Based on scoring system developed for this study and described in Appendix.

Table 5. Mean scores on the Haptic Visual Matching Test for five groups of children during the four phases of testing. (Standard error = 0.65; $\bar{n}_h$ = 21.96)

Phase \ Group	A	B	C	D	E
I	-	-	-	10.89	-
II	9.92	9.14	10.39	11.35	10.30
III	10.20	10.61	12.75	11.65	11.54
IV	11.88	11.54	12.68	12.20	12.71

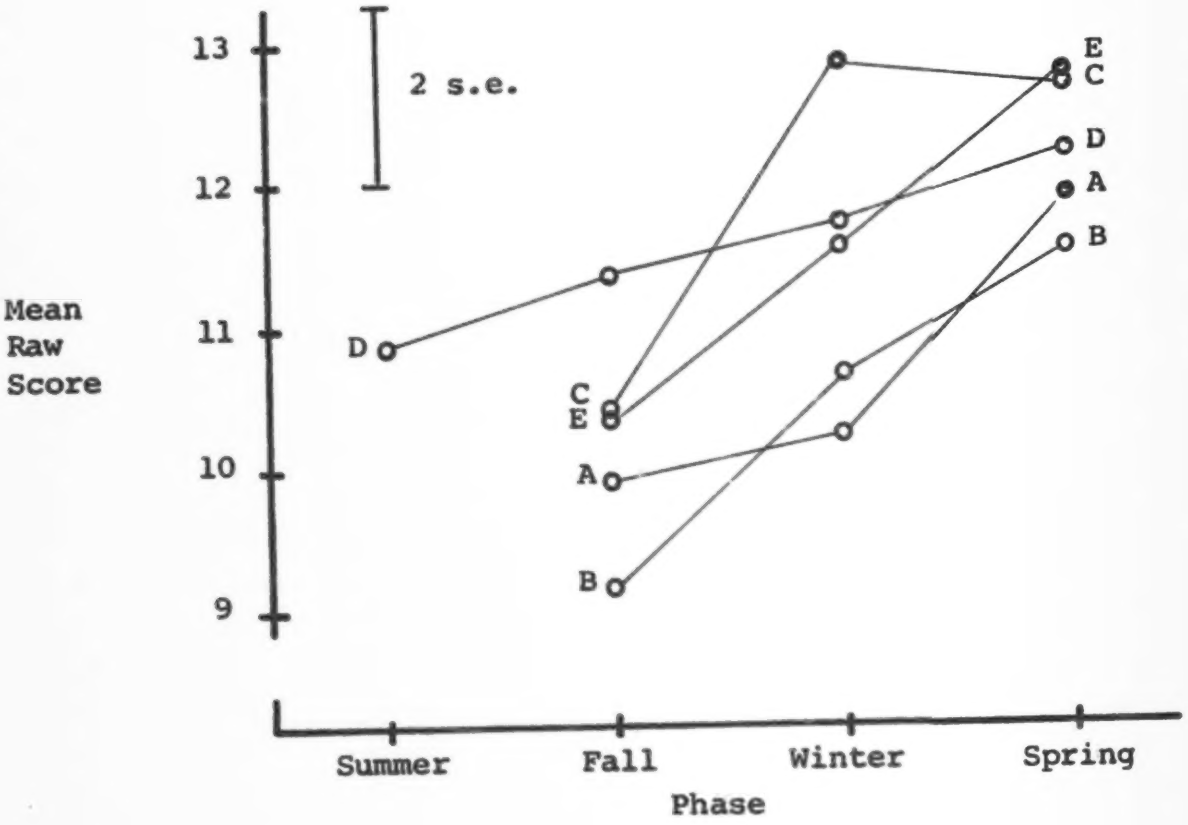

Figure 4. Graphs of mean scores on the Haptic Visual Matching Test for five groups of children during the four phases of testing.

For all groups on all four instruments in this category, the average gain or growth was considerably more than one standard error per phase. Only in case of the Columbia were there norms to facilitate the interpretation of these gains. However, that the gains on the other instruments were measurable in units of standard errors elucidates the point that these are realistic measures for use in measuring growth or evaluating programs for classroom size groups over periods as short as three months.

Category 2: Basic Comprehension Skills

Five of the instruments used have as their primary focus the dimension of comprehension skills. The Picture Vocabulary subtest of the French Pictorial Test of Intelligence (PTI) requires subjects to recall previously acquired verbal meanings as a measure of verbal comprehension. Three subtests of the Illinois Test of Psycholinguistic Abilities (ITPA), Visual-Motor Association, Auditory Decoding and Visual Decoding were chosen to get at other aspects of comprehension. The Visual-Motor Association subtest requires the subject to comprehend relationships among pictures or objects that may go together. The Visual and Auditory Decoding subtests are measures fo the ability to comprehend or obtain meaning from pictures and from the spoken word. A fifth measure in this category is the Symbol Recognition test which requires subjects to differentiate among sets of realistic and abstract symbols.

The mean scores of the five groups on these instruments are reported in Tables 6 through 11 and graphed in Figures 5 through 10.

Table 6. Mean scores on the Picture Vocabulary subtest of the French PTI for five groups of children during four phases of testing. (Standard error = 0.86; $\bar{n}_h = 24.23$)

Phase \ Group	A	B	C	D	E
I	11.80	-	-	15.90	-
II	13.06	10.52	17.71	16.77	18.30
III	15.64	14.13	21.09	18.54	18.96
IV	17.52	13.47	21.63	20.74	21.54

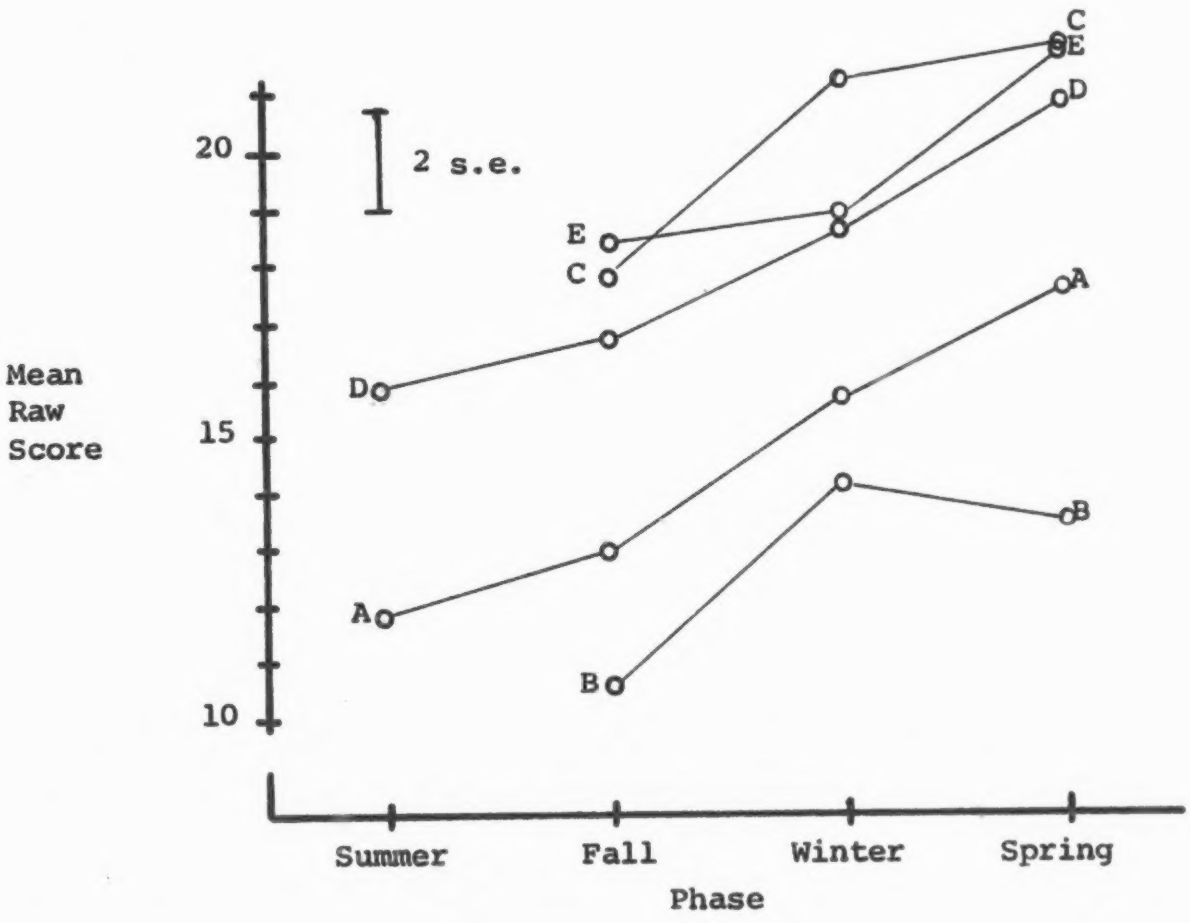

Figure 5. Graphs of mean scores on the Picture Vocabulary subtest of the French PTI for five groups of children during four phases of testing.

Table 7. Mean scores on the Visual-Motor Association subtest of the ITPA for five groups of children during four phases of testing. (Standard error = .89; $\bar{n}_h$ = 23.33)

Phase \ Group	A	B	C	D	E
I	7.57	-	-	9.87	-
II	8.72	7.97	10.93	11.62	10.73
III	10.31	8.74	13.52	11.41	10.27
IV	13.63	13.38	14.38	11.84	14.17

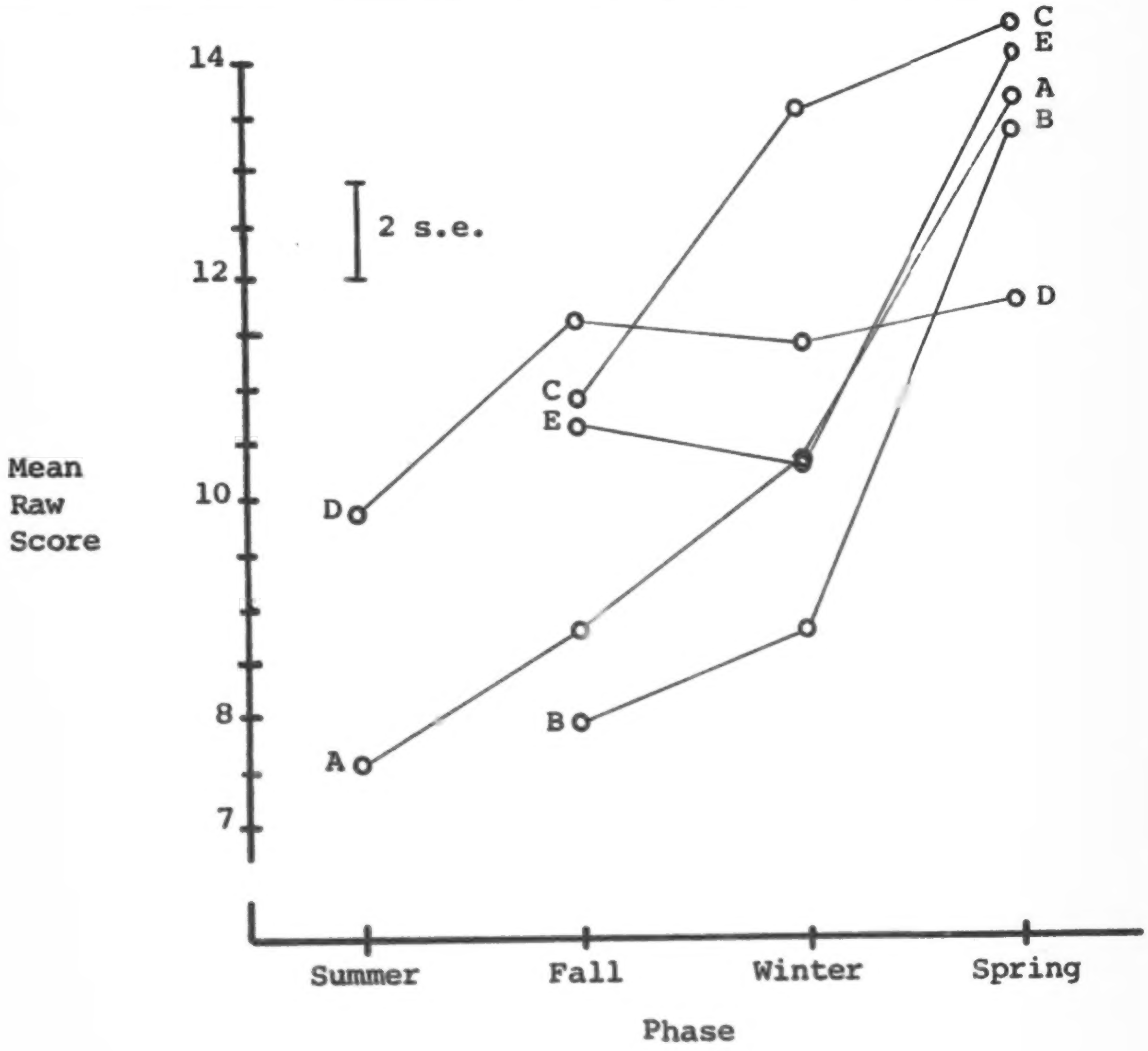

Figure 6. Graphs of mean scores on the Visual-Motor Association subtest of the ITPA for five groups of children during four phases of testing.

Table 8. Mean scores on the Visual Decoding subtest of the ITPA for five groups of children during four phases of testing. (Standard error = .75; $\bar{n}_h$ = 23.33)

Phase \ Group	A	B	C	D	E
I	7.17	-	-	9.17	-
II	8.11	6.62	11.53	11.92	9.90
III	9.73	9.10	12.64	11.13	12.62
IV	11.21	9.46	13.94	12.90	12.61

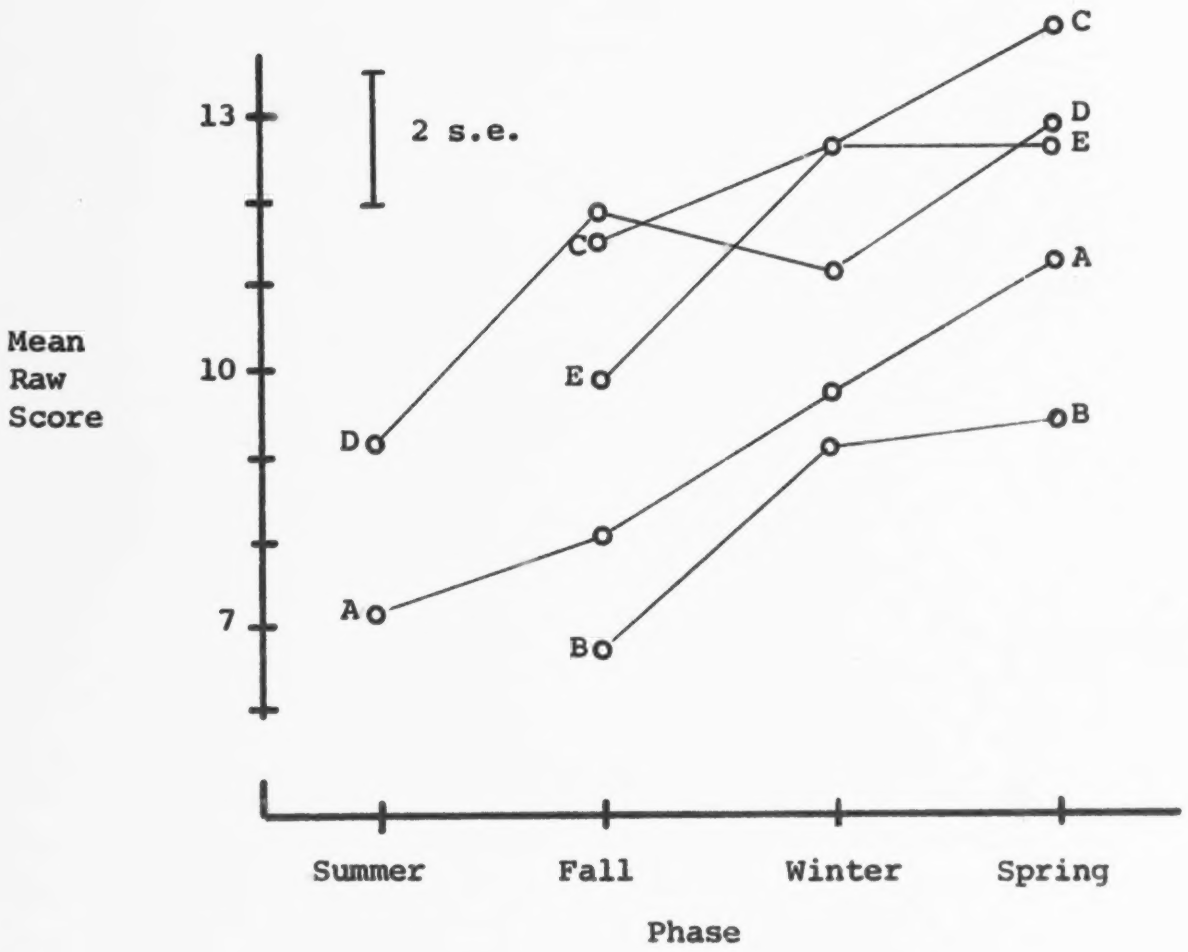

Figure 7. Graphs of mean scores on the Visual Decoding subtest of the ITPA for five groups of children during four phases of testing.

Table 9. Mean scores on the Auditory Decoding subtest of the ITPA for five groups of children during four phases of testing. (Standard error = 1.17; $\bar{n}_h$ = 23.33)

Group Phase	A	B	C	D	E
I	13.47	-	-	17.17	-
II	14.55	12.31	19.32	17.92	20.83
III	15.58	15.53	18.28	20.82	24.12
IV	15.83	16.84	21.56	19.84	25.57

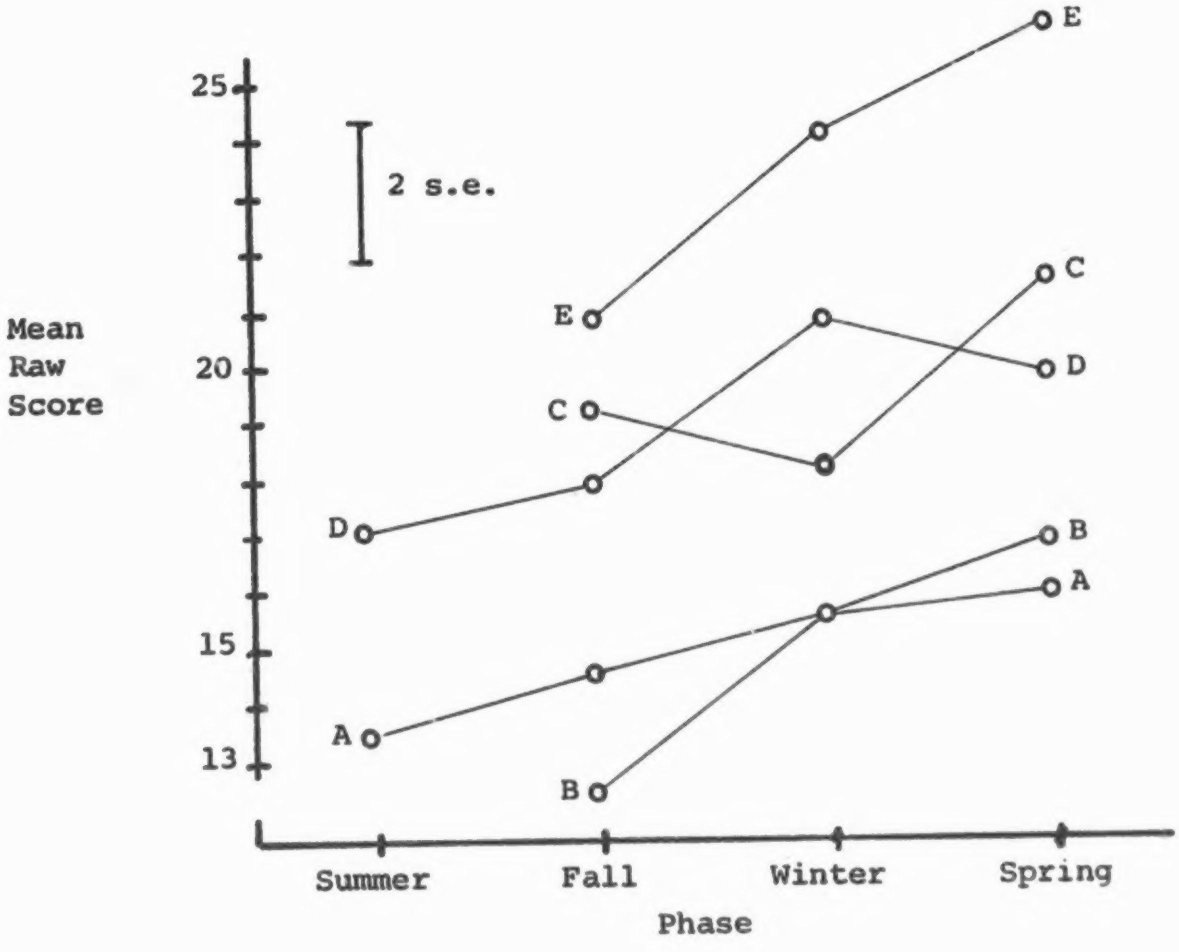

Figure 8. Graphs of mean scores on the Auditory Decoding subtest of the ITPA for five groups of children during four phases of testing.

Table 10. Mean scores on the Symbol Recognition Test, Part I, for five groups of children during four phases of testing. (Standard error = .84; $\bar{n}_h = 26.02$)

Phase \ Group	A	B	C	D	E
I	19.63	-	-	22.40	-
II	21.36	18.46	21.90	24.46	23.77
III	23.64	19.54	24.33	25.62	25.93
IV	23.46	21.19	26.08	26.05	27.23

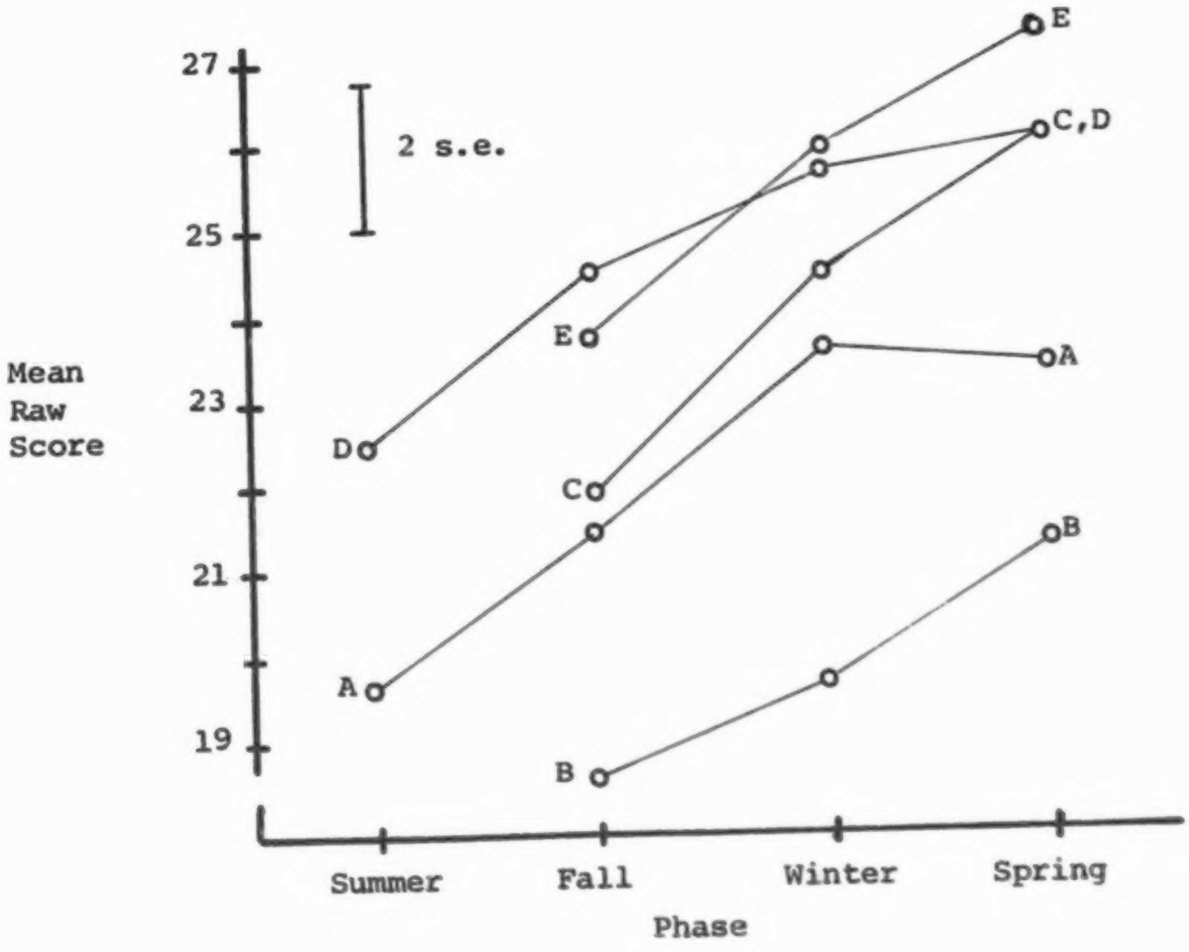

Figure 9. Graphs of mean scores on the Symbol Recognition Test, Part I, for five groups of children during four phases of testing.

Table 11. Mean scores on the Symbol Recognition Test, Part II, for five groups of children during four phases of testing. (Standard error = .46; $\bar{n}_h$ = 26.02)

Phase \ Group	A	B	C	D	E
I	26.47	-	-	28.50	-
II	27.67	26.32	29.00	29.11	29.00
III	27.96	26.54	29.70	29.48	29.56
IV	28.75	28.14	30.20	29.95	29.86

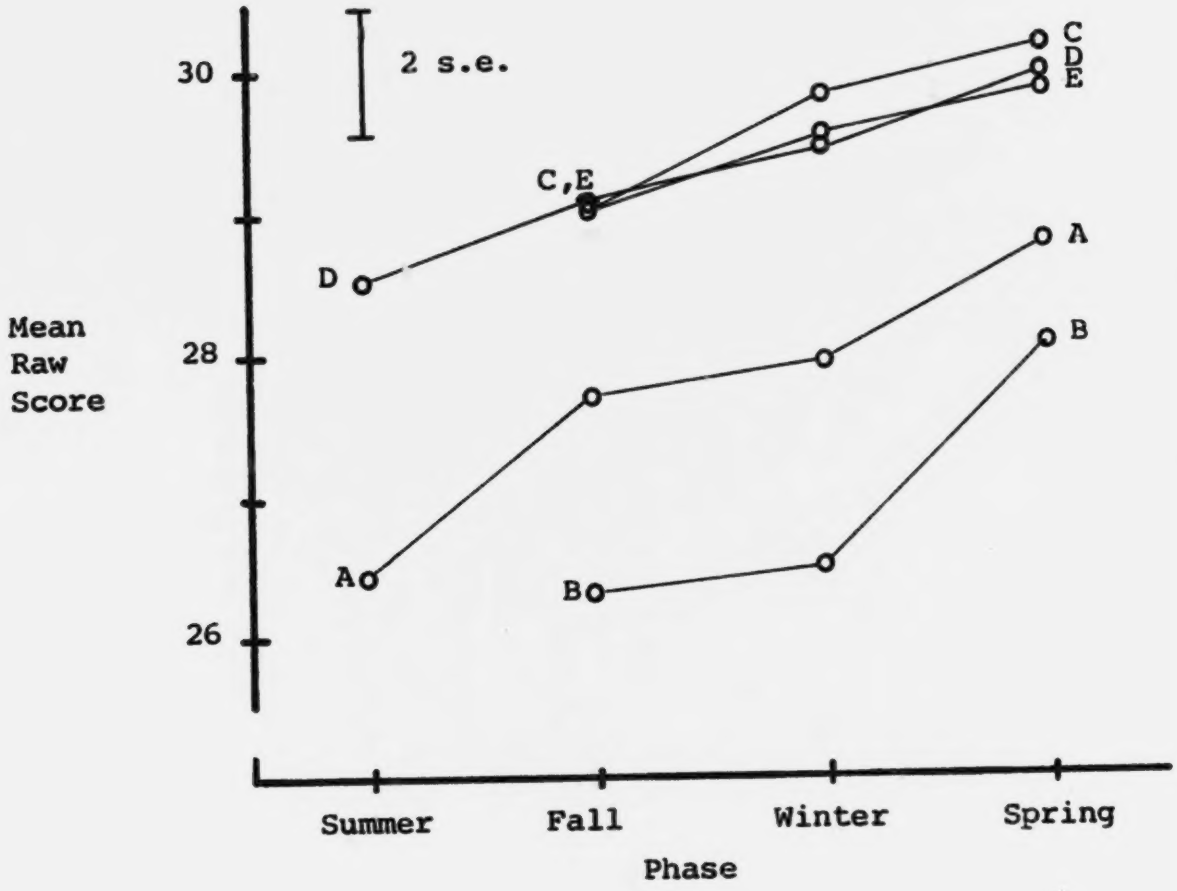

Figure 10. Graphs of mean scores on the Symbol Recognition Test, Part II, for five groups of children during four phases of testing.

Results of the French PTI Picture Vocabulary subtest are presented in Table 6 and Figure 5. Again Groups C, D and E are clustered with the means within a two standard error interval at both Phase II and Phase IV and within three standard errors at Phase III. The curve for Group A is three to four standard errors below this cluster and the curve for B is two or more standard errors below that for A. As with the instruments in the first category, all five groups increased in performance across time with the increases for all groups between Phase II and Phase IV being equivalent to a gain of at least 12 months in mental age.*

The Visual-Motor Association subtest of the ITPA (Table 7 and Figure 6) provides one of the few tasks on which Groups A and B performed as well as the others. Although the initial performance of these two groups was substantially lower than those of the other three, the mean gains of over three standard errors between the winter and spring testings brought A and B to par with Groups C and E. Group D changed very little on this task and at Phase IV was at least two standard errors below A, C and E.

Table 8 and Figure 7 summarize the Visual Decoding subtest of the ITPA. Again, one notes the pattern of C, D and E forming a cluster of generally higher performance with Group C performing best.

The third subtest of the ITPA, Auditory Decoding, is summarized in Table 9 and Figure 8. The five groups divide into three clusters on this measure with the two inner core groups, Groups A and B, averaging approximately three standard errors less than Groups C and D which in turn average over three standard errors below Group E.

---

*Based on norms in the Manual, Pictorial Test of Intelligence.

Results of the final instrument in this subset, the Symbol Recognition Test, appears in Tables 10 and 11 and Figures 9 and 10. Part I of this test requires the subject to choose the appropriate abstract symbolization of a familiar object while Part II requires the subject to correctly identify a realistic symbolization of the familiar object. Thus, the scores on Part II provide an upper bound for the scores in Part I. The maximum score on either part is 31, so the Part II scores are very close to the ceiling, particularly with Groups C, D and E. Subtracting the Part I mean scores from the Part II mean scores for Phase IV results in differences for Groups A through E of 5.29, 7.05, 4.12, 3.90 and 2.63. Thus, not only do Groups A and B do less well on the two parts, but there is also a greater discrepancy between the two parts indicating more difficulty in handling symbols as well as a smaller repertoire of visual concepts.

In summary, those tests chosen to measure basic comprehension skills presents the same general picture as the discrimination tasks in Category 1, i.e., with Groups C, D and E outperforming the inner city disadvantaged groups, A and B. Within the cluster of C, D and E it is interesting to note that only on the auditory task does Group E, the only clearly middle class group, do substantially better.

Category 3: Basic Quantitative Skills

The two measures in this category are the Size and Number subtest of the French PTI and the Elkind Measurement of Quantitative

Table 12. Mean scores on the Size and Number subtest of the French PTI for five groups of children during four phases of testing. (Standard error = .86; $\bar{n}_h$ = 24.23)

Group Phase	A	B	C	D	E
I	5.93	-	-	10.10	-
II	7.22	6.41	9.21	11.39	12.07
III	7.57	7.88	12.48	14.04	15.89
IV	10.28	7.89	13.90	15.42	17.58

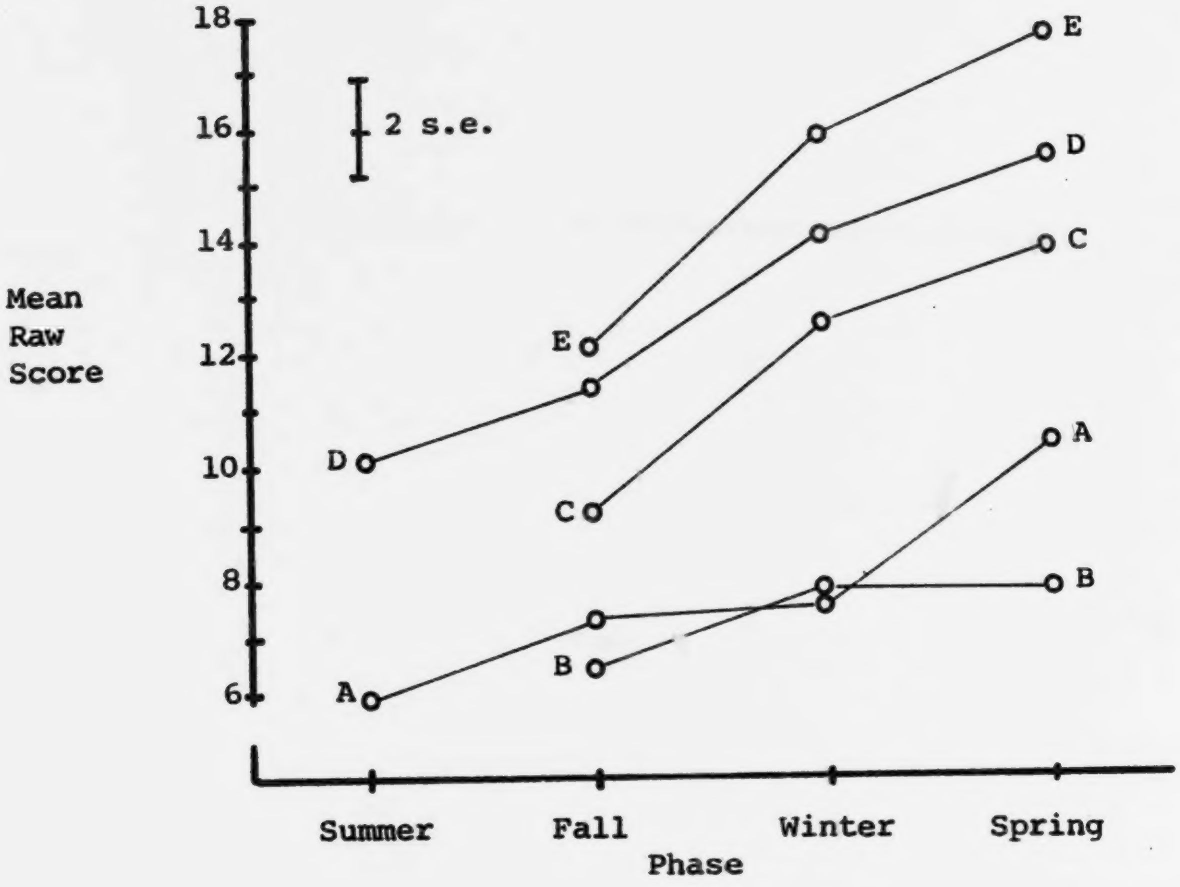

Figure 11. Graphs of mean scores on the Size and Number subtest of the French PTI for five groups of children during four phases of testing.

Comparisons instrument. The Size and Number subtest samples the subject's perception and recognition of size, number symbol recognition and comprehension, ability to count and ability to solve simple arithmetic problems. The Elkind procedure provides an objective measure of the Piagetian ideas of conservation of number and of volume. The measure was scored so as to produce a total score on quantitative comparisons of number and a total score on volume comparisons. Three extensive comparisons[*] are made and proportions of subjects in each group making 1 or 2 correct responses in two trials were recorded.

Table 12 and Figure 11 record the results on the French PTI Size and Number subtest. Groups C, D and E performed measurably better with great similarity in the shapes of their curves. With the exception of the D-C difference, all Phase IV differences are of at least two standard errors in magnitude indicating increasing differentiation over time. The manual for the French PTI reports mental ages corresponding to subtest scores. A raw score of 6 corresponds to a MA of 4 years, 0 months, a raw score of 9 to 4-6, a raw score of 11 to 5-0, and a raw score of 16 to 6-0. Thus, the mean performance of the A and B groups at Phase IV is still below the norm for 5 years and only Group E has a mean above the 6 year norm.

The Elkind results for the five groups are reported in Tables 13 and 14 and Figure 12. Table 13 presents the mean number of correct responses to sequences of questions designed to measure the

---

[*] See Appendix

Table 13. Mean scores on the Elkind Measurement of Quantitative Comparisons for five groups of children given during the spring semester of kindergarten. (Maximum score = 12)

	Groups				
Score	A	B	C	D	E
Conservation of number	9.73	9.64	10.57	9.50	11.00
Convservation of volume	7.86	8.64	8.14	7.56	8.62

Table 14. Proportions of subjects in five groups responding correctly to three types of extensive comparison items on the Elkind Measurement of Quantitative Comparisons instrument.

Item	Correct Responses	s_p	Group A	B	C	D	E
Quantity	1	.10	.23	.21	.14	.22	.14
	2	.11	.41	.36	.43	.39	.67
Volume (into 2 parts)	1	.11	.10	.13	.28	.27	.14
	2	.10	.32	.29	.24	.17	.29
Volume (shape)	1	.08	.05	.07	.14	.17	.19
	2	.10	.18	.29	.38	.11	.38

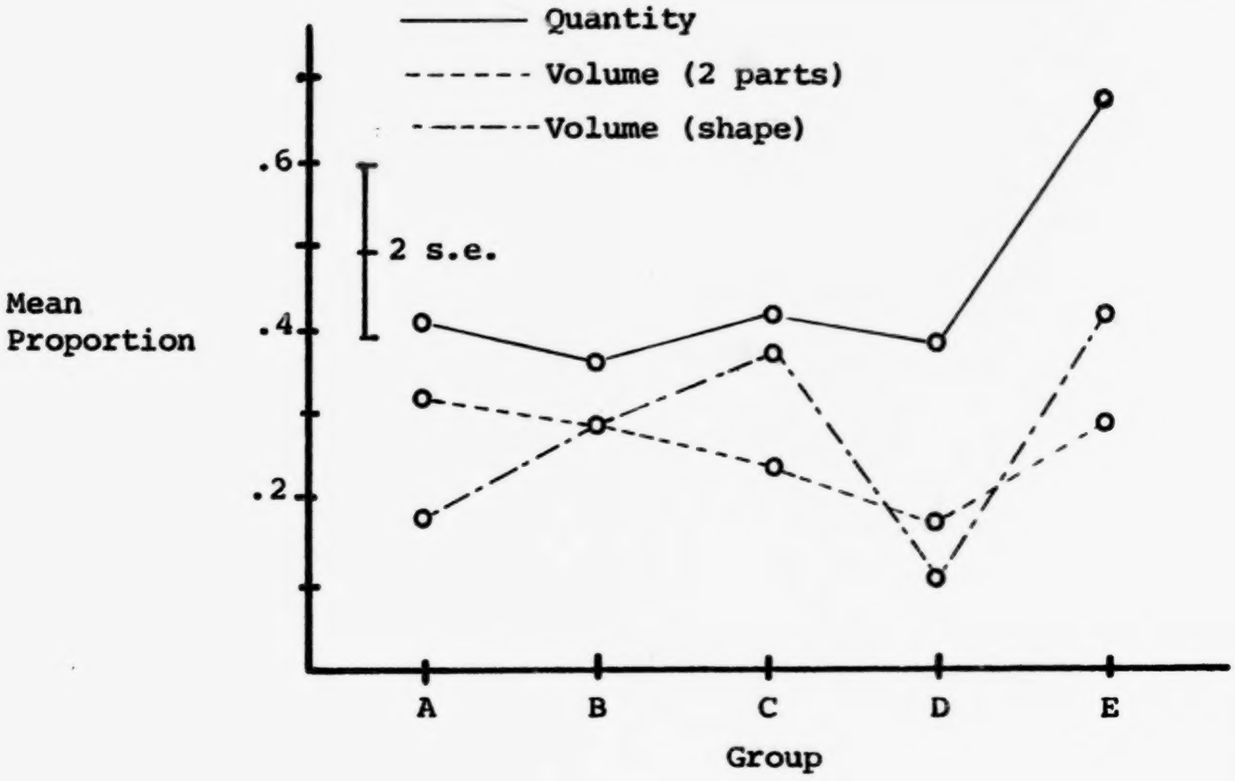

Figure 12. Graphs of proportions of children in five groups successfully completing three types of conservation tasks on the Elkind Measurement of Quantitative Comparisons.

extent to which subjects have attained conservation of number and of volume. Group E has the highest mean on the number sequence and Group B has the highest on the volume sequence. Table 14 reports the proportions of subjects in each group answering the three pairs of extensive comparison questions correctly. The proportion answering the two items in any set correctly may be assumed to have the conservation skill. Note that two-thirds of Group E have achieved conservation of number, but that the best results on the volume tasks are little better than one-third. While the data has been reported in both ways note that the variations in Table 13 are almost entirely a function of the extensive items reported in Table 14. Figure 12 presents those proportions of subjects answering both items in any pair correctly.

Category 4: General Ability and Knowledge

Four of the instruments used are in this category. They include three widely-used standardized instruments, the Pintner-Cunningham Primary Test, the Caldwell Preschool Inventory and the Wechsler Pre-School and Primary Scale of Intelligence (WPPSI). Because of the length and the relatively global nature of these instruments, they were only administered once, the Caldwell during Phase III and the Pintner-Cunningham and WPPSI during Phase IV. The fourth instrument, administered during all Phases, was a General Information Test designed to measure a child's general knowledge of his environment.

Table 15 and Figure 13 present the results on the General Information Test. Figure 13 shows the same general clustering pattern

Table 15. Mean scores on the General Information Test for five groups of children during four phases of testing. (Standard error = 1.13; $\bar{n}_h$ = 24.09)

Phase \ Group	A	B	C	D	E
I	12.23	-	-	19.83	-
II	15.75	13.07	19.89	20.65	23.43
III	17.46	14.44	22.65	23.36	26.51
IV	19.84	17.16	26.21	27.18	28.96

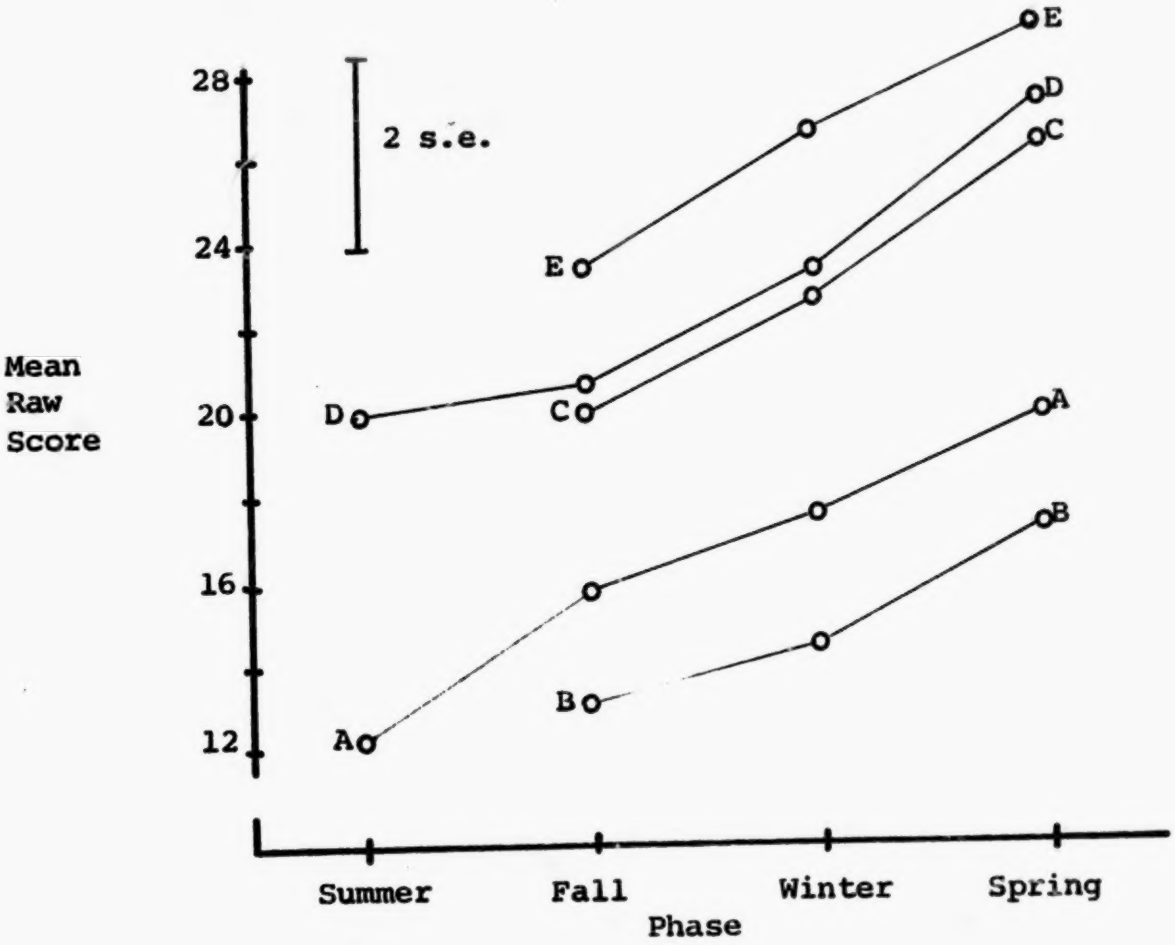

Figure 13. Graphs of mean scores on the General Information Test for five groups of children during four phases of testing.

with Groups C, D and E performing considerably better at all stages than Groups A and B. Comparison to the standard of two standard errors reveals that all differences between groups except between Groups D and E and between C and D are at least this large.

Table 16. Mean raw scores and intelligence quotients on the Pintner-Cunningham Primary Test for five groups of children tested during the spring or fourth phase of testing. (Standard error of mean IQ = 2.89; $\bar{n}_h = 21.07$)

Mean of:	Group A	B	C	D	E
Raw scores	32.7	38.0	42.0	49.7	58.7
IQ scores	86.0	88.2	91.9	94.8	103.0

The Pintner-Cunningham results appear in Table 16 in both mean raw score and mean I.Q. score form. Only Group E is at or close to age norms as indicated by the I.Q. scores. Groups C and D have mean I.Q. scores in the low 90's and Groups A and B have mean I.Q. scores in the high 80's. The Pintner-Cunningham is regularly administered to kindergarten children in the city of Milwaukee and provides part of the criterion data upon which promotion to first grade is based. The mean of the three urban groups is over ten I.Q. points below the theoretical mean of 100.

The Caldwell Preschool Inventory, results of which are tabled in Table 17 and graphed in Figure 14, provides four part scores and a total score. The part scores are identified as measuring personal-

Table 17. Mean scores on the Caldwell Preschool Inventory for five groups of children during the winter of kindergarten year. (Standard error for total = 1.87; $\bar{n}_h$ = 27.98)

Subtest	Group A	B	C	D	E
1. Personal-social responsiveness	18.7	16.2	21.1	21.0	22.1
2. Associative vocabulary	10.2	8.6	15.7	14.9	16.8
3. Concept activation-numerical	9.5	8.8	12.1	14.0	15.3
4. Concept activation-sensory	13.1	11.9	16.7	17.1	18.0
Total	51.5	45.4	65.6	67.0	72.2

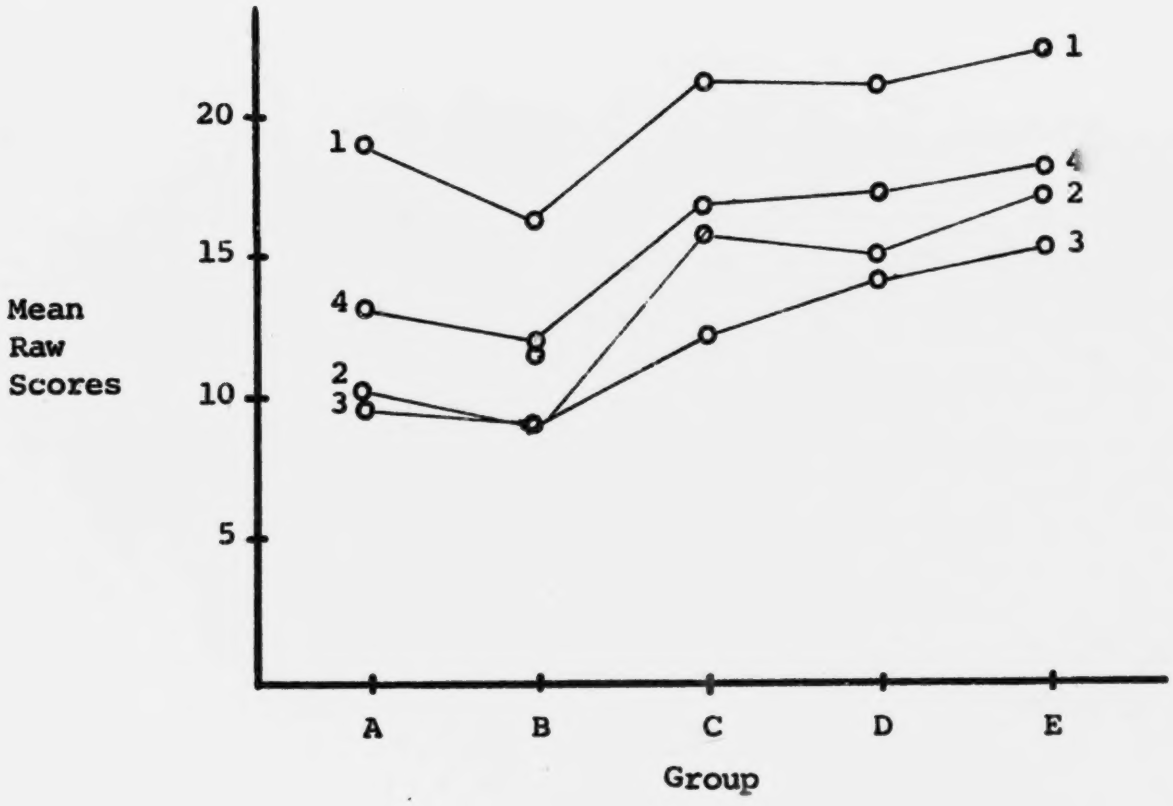

Figure 14. Graphs of score profiles of five groups on four subtests of the Caldwell Preschool Inventory. (Numbering of subtests corresponds to that in Table 17.)

social responsiveness, associative vocabulary, concept activation-numerical and concept activation-sensory. In terms of overall performance the familiar clustering of C, D and E and A and B appears (see Figure 14). Of more interest is the striking similarity in the general factor underlying all four of the subtests.

The final general ability test is the WPBSI which was administered during the final weeks of the kindergarten year. Mean raw scores and standard deviations on each of the ten subtests for each of the five groups are presented in Table 18. The mean scale scores for these subtests as well as the Total, Verbal and Performance scale scores are presented in Table 19 and Figures 15, 16 and 17.

Relative performance across all ten subtests was consistent with Group E performing best, followed by Groups C and D with Groups A and B at least a full scale point lower on all except the Block Design subtest. The regularities of this clustering are especially pronounced on the Verbal tests (see Figure 15). Groups A, C, D and E tended to do slightly better on the Information and Arithmetic subtests than on the other Verbal tests. The best Verbal performance for Group B was on the Similarities subtest.

Results of the Performance subtests, graphed in Figure 16, are somewhat less regular than on the Verbal, although the scores of Group E are consistently the highest followed by those of Groups C and D and then Groups A and B. The Mazes subtest was particularly difficult for the urban disadvantaged groups, resulting in mean scale scores less than 7.00. The only mean above 11.00 is the 12.10

Table 18. Mean raw scores and standard deviations on ten subtests of the WPPSI for each of five subgroups.

	Group									
Subtest	A(N=34)		B(N=20)		C(N=26)		D(N=26)		E(N=29)	
	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.	$\bar{X}$	s.d.
Verbal										
Information	11.71	2.82	10.05	4.11	13.58	2.53	14.58	2.53	15.52	2.23
Vocabulary	11.65	4.47	10.85	5.49	16.89	6.98	16.12	5.79	19.83	4.88
Arithmetic	9.47	2.14	8.15	3.87	11.23	2.41	12.62	2.89	13.17	2.30
Similarities	7.24	3.36	7.35	4.07	9.62	3.03	10.12	3.59	11.76	4.11
Comprehension	11.06	4.59	9.55	5.77	14.58	5.01	15.04	5.37	16.07	4.17
Performance										
Animal House	36.53	11.58	29.40	16.55	46.23	7.67	47.00	13.37	49.48	11.00
Picture Completion	10.91	4.16	9.90	3.92	12.81	3.27	15.27	8.17	15.35	3.46
Mazes	6.82	3.37	7.35	4.94	14.31	5.84	14.39	5.67	16.69	6.85
Geometric Design	8.53	4.97	7.55	5.30	12.35	8.24	12.54	6.66	16.31	5.48
Block Design	9.35	4.28	9.80	5.04	10.73	4.67	13.42	3.41	14.07	5.29

Table 19. Mean scale scores on ten subtests and the Verbal, Performance and Total scales of the WPPSI for each of five subgroups.

Subtest or Scale	Group A(N=34)	B(N=20)	C(N=26)	D(N=26)	E(N=29)
Verbal					
Information	7.97	6.75	9.65	9.85	10.72
Vocabulary	6.97	6.30	9.23	8.50	9.90
Arithmetic	7.94	6.65	9.62	9.88	10.31
Similarities	7.12	7.40	9.08	8.62	9.79
Comprehension	7.44	6.80	9.42	9.04	9.66
Performance					
Animal House	7.65	6.60	9.58	9.77	9.97
Picture Completion	8.47	7.65	9.58	9.42	10.62
Mazes	6.79	6.90	10.15	9.27	10.66
Geometric Design	8.18	7.55	9.85	9.62	12.10
Block Design	7.94	8.50	8.65	9.85	10.24
Verbal	37.85	34.40	47.39	46.12	50.93
Performance	39.82	37.60	48.08	47.69	53.07
Total	77.82	72.15	95.31	94.39	103.94

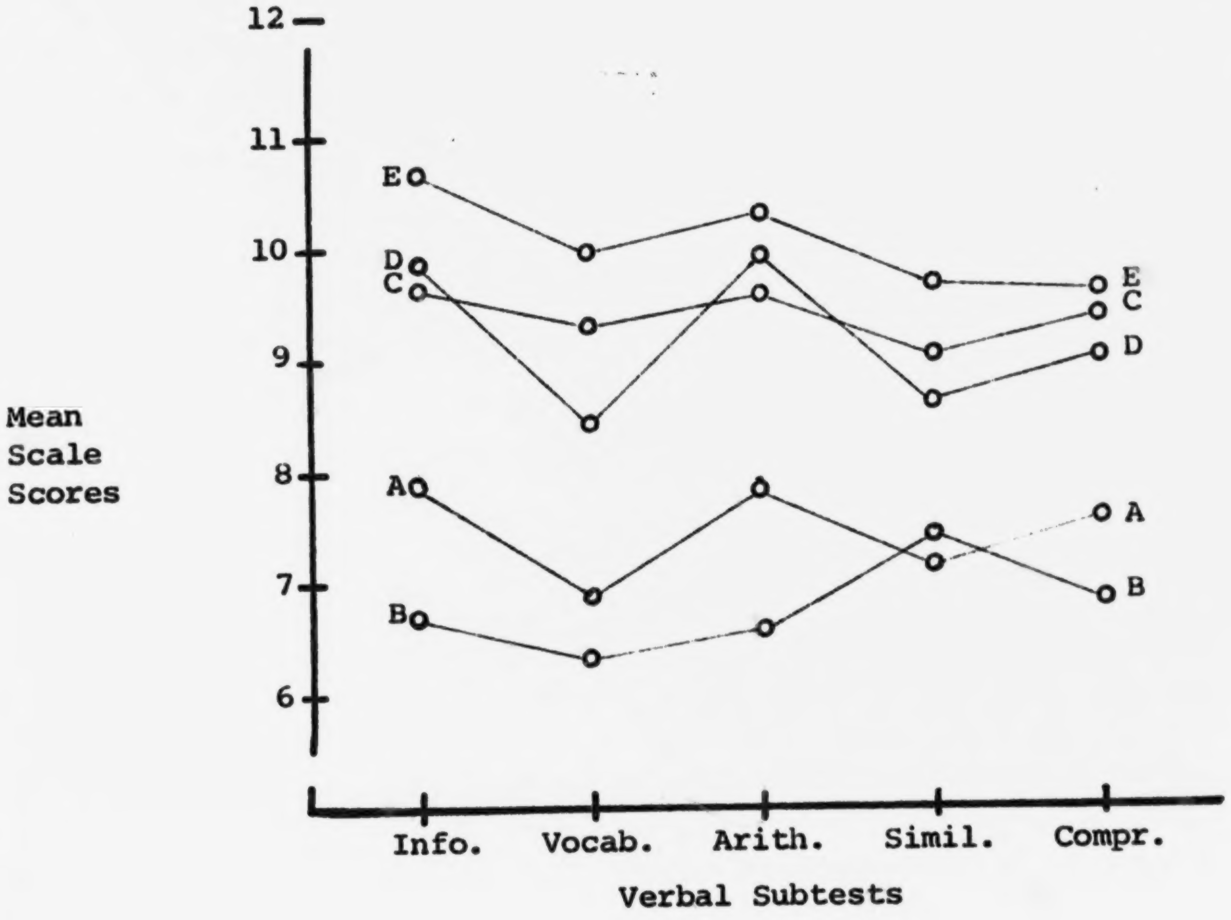

Figure 15. Graphs of mean scale scores on the five Verbal subtests of the WPPSI for five groups of children. (Administered during final month of kindergarten.)

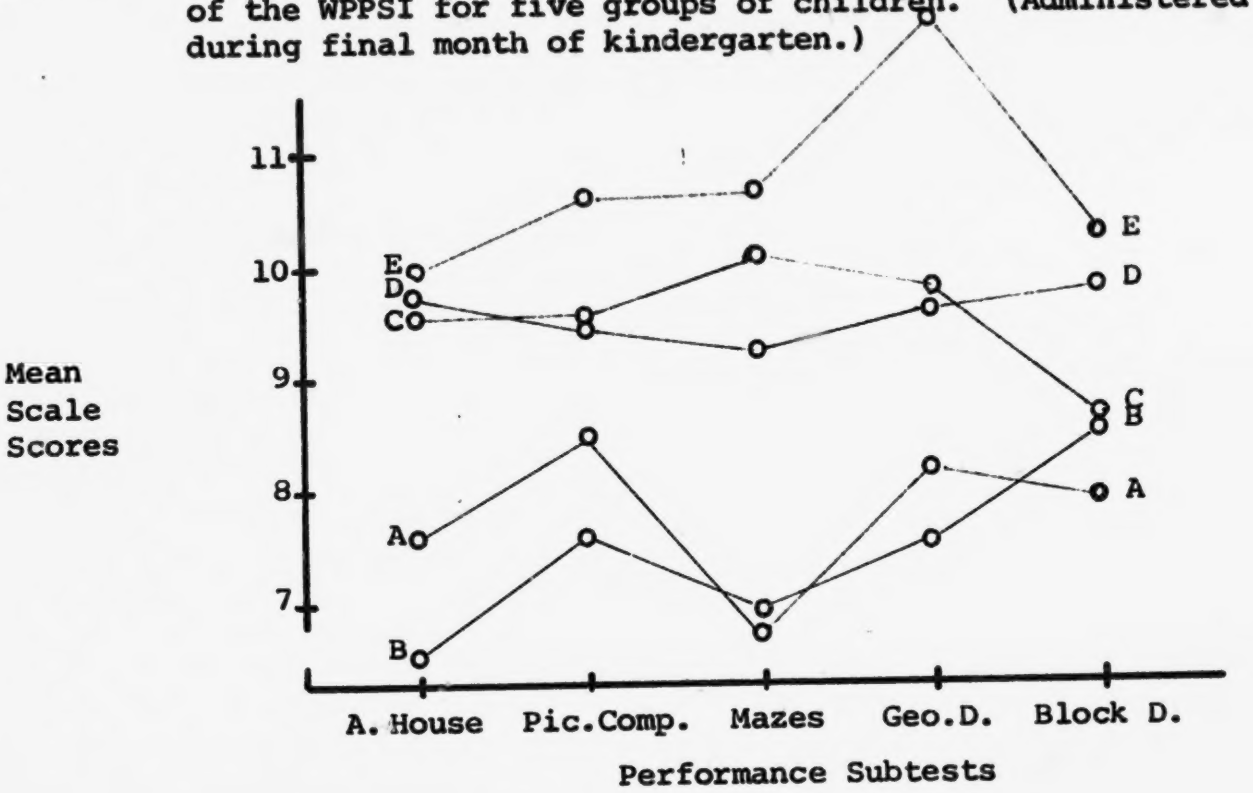

Figure 16. Graphs of mean scale scores on the five Performance subtests of the WPPSI for five groups of children. (Administered during final month of kindergarten.)

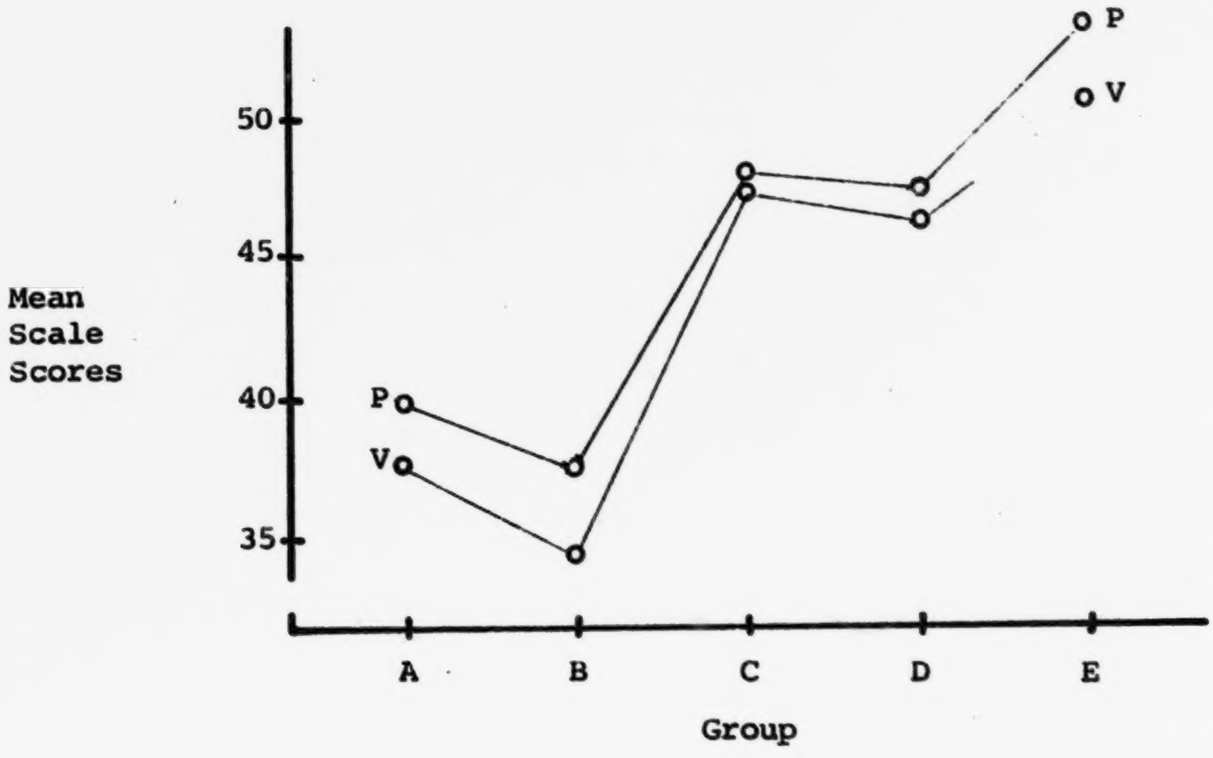

Figure 17. Graphs of mean total scale scores on the Verbal and Performance Scales of the WPPSI for five groups of children.

attained by Group E on the Geometric Design.

Mean total scores for the Performance and Verbal Scales are graphed in Figure 17. Verbal mean scores are consistently 1.5 to 2.5 scale points lower than the Performance scores. The clustering apparent in the subtest scores is again seen with the means for A and B between 34 and 40, those for C and D between 46 and 49 and those for E over 50.

Category 5: Cognitive Style

The two instruments in this category were designed to measure aspects of how children approach cognitive tasks rather than of how well they perform. One instrument, the Haptic Visual Matching Test, was discussed earlier under Category 1. This test was also timed and the times from presentation of stimulus to response was taken as a measure of an impulsivity-reflectivity dimension. A second instrument, the Matching Familiar Figures Test, was also designed to measure impulsivity-reflectivity. However, analysis of response times from Phases I and II indicated that there was no useful information in that measure and it will not be reported. The other aspect of cognitive style that we attempted to measure was the classification structure that is used in approaching a cognitive task. The Block Sorting Test attempted to differentiate among groups as to preferences for categorizing by color, size or shape.

Table 21. Mean response times (in seconds) on the Haptic Visual Matching Test for five groups of children during four phases of testing. (Standard error = 14.1; $\bar{n}_h$ = 21.96)

Group Phase	A	B	C	D	E
I	-	-	-	96.3	-
II	12.70	93.0	104.0	97.8	89.4
III	77.4	82.0	92.4	85.7	92.4
IV	99.9	88.5	96.7	100.6	100.5

Table 21 presents the response times on the Haptic test. While there was considerable variability within groups, the groups themselves did not differ greatly on this measure. Excluding the mean of 127.0 seconds by Group A in Phase II, all other means are within two standard errors of the lowest mean, 77.4.

The results of the Block Sorting Task, which was only administered once during either Phase III or IV, are found in Table 22 and

Table 22. Mean frequencies of categorizations on basis of color, size or shape on the Block Sorting Task for five groups of children.

Group	2-Dimensional Stimulus Color	 Size	 Shape	3-Dimensional Stimulus Color	 Size	 Shape
A	7.50	.54	7.62	7.08	.96	7.46
B	10.00	1.15	4.62	10.50	1.00	4.07
C	4.42	.84	10.74	4.79	.21	11.32
D	2.82	.47	12.65	1.88	.41	13.71
E	3.29	.14	12.57	3.00	.10	12.86

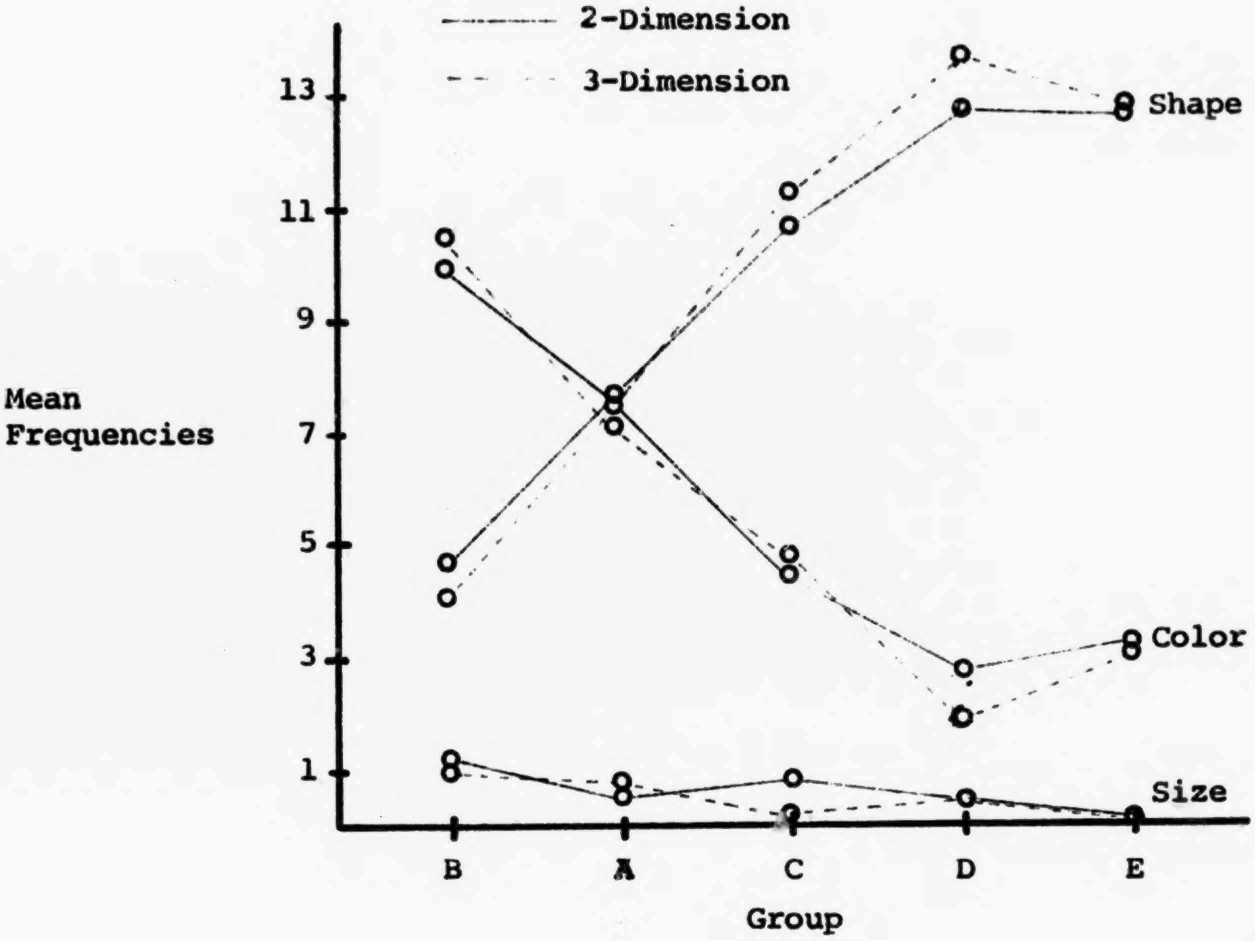

Figure 18. Mean frequencies of color, size and shape categorizations on Block Sorting Task by five groups of children.

Figure 18. Both 2- and 3-dimensional stimuli were used with unusually consistent results. As will be seen by examination of Figure 18, size was a relatively unimportant dimension relative to the other two. Since the three classification dimensions are related, one finds an inverse relationship between the color and shape dimensions that clearly differentiates among the groups. Group B made over twice as many categorizations on the basis of color than on shape. Group A was about equally divided in frequency of color and shape categorizations. The results of Group C are almost the opposite of those for B with twice as many shape decisions and the children in Groups D and E preferred to categorize by shape by a ratio of over 4 to 1.

Category 6: Affective Domain

Two techniques were used in this category to obtain a measure of the subjects' self-concept and to assess propensity toward immediate gratification. A modification of the Illinois Index of Self-Derrogation was used to elicit self-estimates from children. Ability to delay gratification was examined by a procedure in which the child was offered a choice between an immediate reinforcement or a delayed, but more desirable, reinforcement. This technique is referred to by Delay of Gratification.

The self-concept scores may be presented in either positive or negative form. Table 23 presents the results in terms of positive responses. These are graphed in Figure 19. The two non-disadvantaged Groups, C and E, have clearly more positive self-concepts than the urban disadvantaged groups. While Group D is closer to Groups A and B

Table 23. Mean scores on the Illinois Index of Self Derrogation for five groups of children during four phases of testing. (Standard error = 1.05; $\bar{n}_h = 20.79$)

Phase \ Group	A	B	C	D	E
I	17.79	-	-	-	-
II	18.89	18.50	22.89	20.69	23.20
III	18.82	19.55	23.17	-	-
IV	20.47	20.48	24.88	23.90	25.42

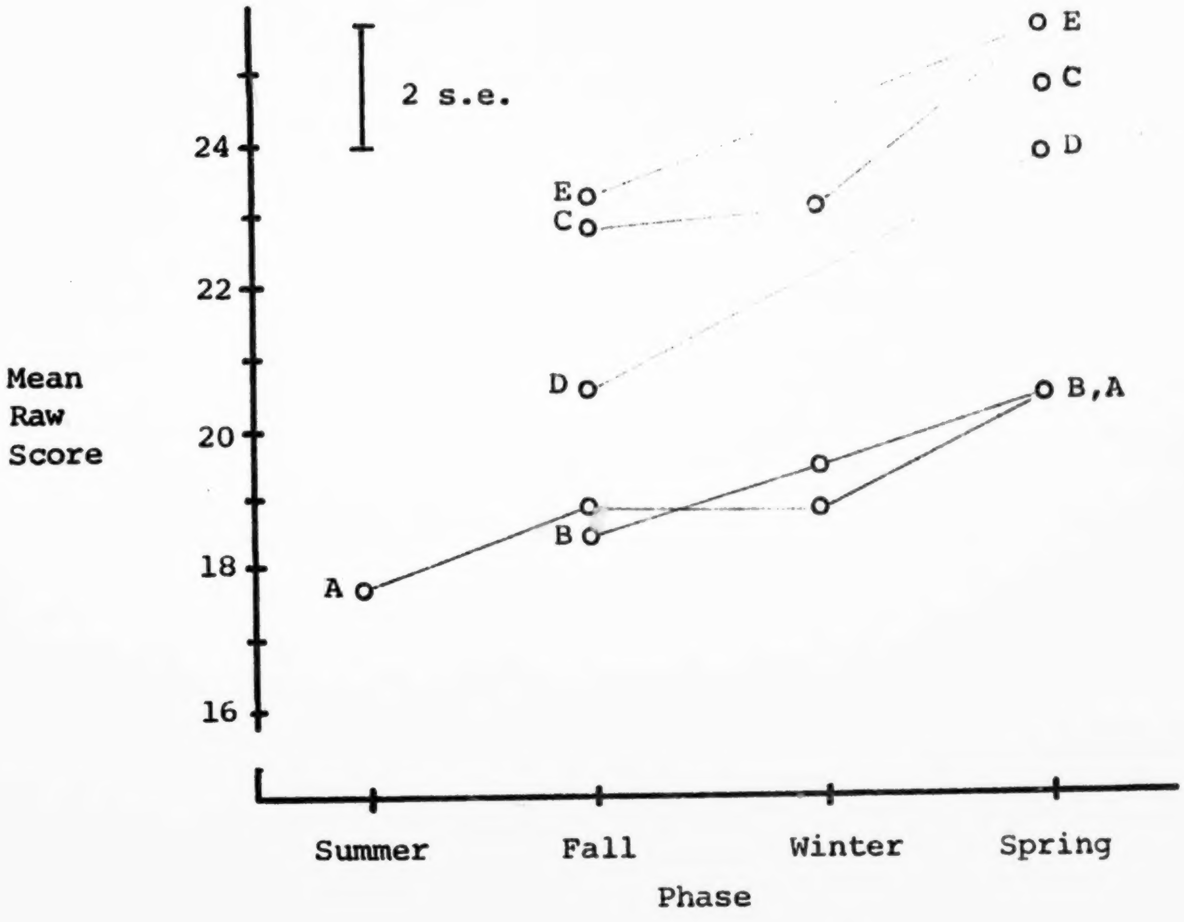

Figure 19. Graphs of mean scores on the Illinois Index of Self Derrogation for five groups of children during four phases of testing.

Table 24. Proportions of groups choosing the delayed reinforcement on the Delay of Gratification measure. For five groups of children during three phases of testing. (Standard error = .09; $\bar{n}_h$ = 25.85)

Phase	Group	A	B	C	D	E
I		.20	-	-	.18	-
II		.26	.36	.50	.15	.17
III		.35	.38	.67	.19	.33
IV	Not administered					

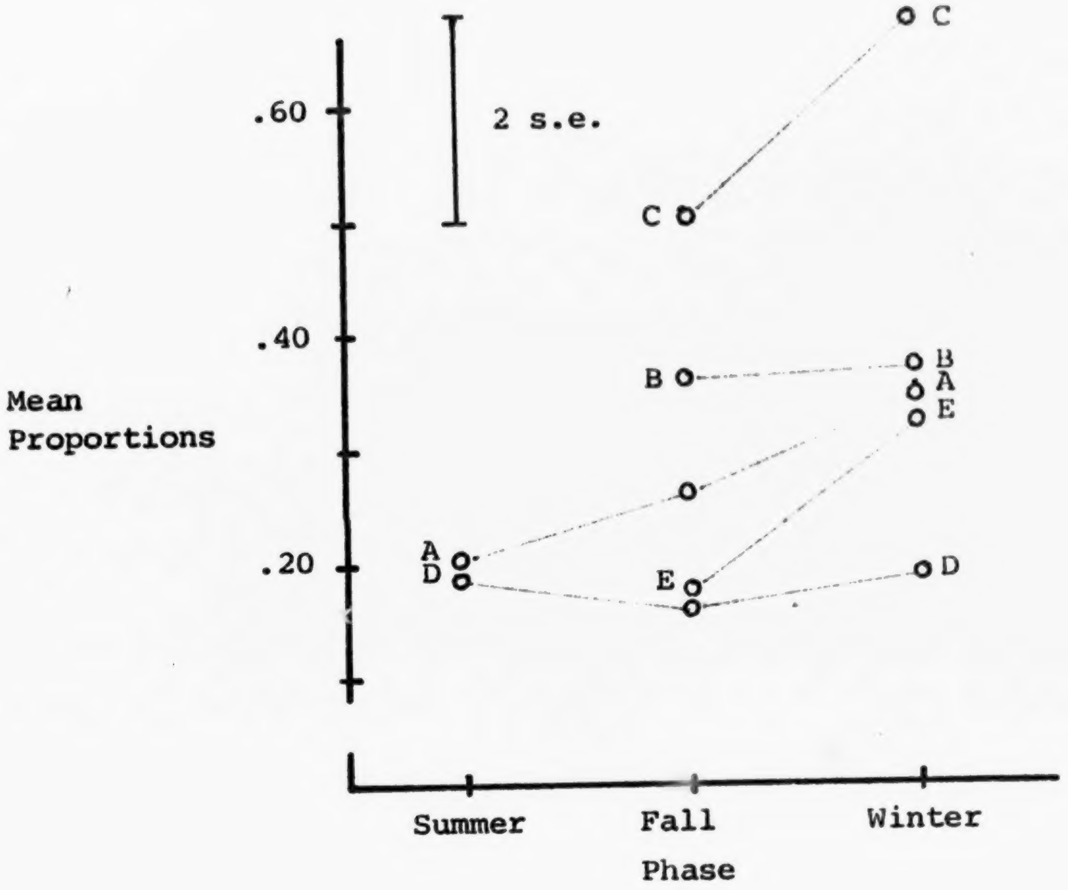

Figure 20. Graphs of proprotions choosing the delayed reinforcement on the Delay of Gratification measure for five groups of children during three phases of testing.

at the initial testing, the results after six additional months of school put the group much closer to C and E.

The Delay of Gratification measure reported in Table 24 and Figure 20 was not given during Phase IV. The data is reported as proportions of children in each group choosing the delayed, and more desirable, reinforcement. At all phases, the small town groups, D and E, manifested the least willingness to delay gratification while the urban, black, non-disadvantaged group, C, delayed over half of the time.

Category 7: Family Relationships

The final category in which differentiation among the five groups was made was in the area of affective responses to individuals in the child's family. The instrument used was the Family Relations Test (FRT) which is summarized in Tables 25 and 26. The FRT allows the child to structure his family for the examiner and then items are presented designed to elicit expressions of feeling by the child toward the several persons in the family. For this study, feelings were categorized as:

Outgoing positive
Outgoing negative
Incoming positive
Incoming negative
Dependency on others

These expressions of feeling were tabulated as being directed toward or received from the child's father, mother, siblings, self or nobody.

Table 25. Mean frequencies of responses attributed to five members of family in five feeling categories on the Family Relations Test. Data for five groups from Phase I or II administration.

	Group				
Category and Person	A	B	C	D	E
Outgoing Positive					
Nobody	.38	.30	.15	.62	.17
Self	1.03	1.74	1.56	1.00	.57
Father	.79	1.19	1.33	1.52	2.00
Mother	2.17	1.44	1.33	2.24	1.93
Siblings	3.79	2.81	3.52	2.86	2.93
Outgoing Negative					
Nobody	1.17	1.00	1.67	2.59	2.23
Self	.90	.81	1.07	.52	.50
Father	.52	.48	1.11	.62	.43
Mother	.66	.41	.44	.72	.70
Siblings	4.03	4.59	3.63	3.59	4.23
Incoming Positive					
Nobody	.41	.63	.15	.72	.10
Self	.93	.81	.67	.62	.43
Father	.75	.96	1.89	1.41	1.67
Mother	1.57	1.56	2.15	2.45	3.50
Siblings	4.14	3.33	3.59	3.24	2.33
Incoming Negative					
Nobody	.93	1.41	1.30	2.66	1.77
Self	.76	1.04	.81	.55	.10
Father	.41	.74	1.11	.72	1.03
Mother	.86	.81	.96	1.10	1.87
Siblings	4.34	3.74	3.74	2.66	3.23
Dependency					
Nobody	.55	.52	.19	.55	.40
Self	.93	.89	.70	.52	.33
Father	.76	1.19	2.04	1.41	1.83
Mother	2.21	2.59	2.67	4.07	4.23
Siblings	2.83	2.59	2.37	1.34	1.10

Table 26. Mean frequencies of responses attributed to five members of family in five feeling categories on the Family Relations Test. Data for five groups from Phase IV administration.

			Group		
Category and Person	A	B	C	D	E
Outgoing Positive					
Nobody	.76	.62	.12	.62	.42
Self	.85	1.24	.81	1.05	.29
Father	1.06	1.10	2.15	1.04	2.54
Mother	2.06	1.19	2.38	1.86	1.88
Siblings	2.79	3.33	3.46	2.62	3.08
Outgoing Negative					
Nobody	1.50	1.76	2.73	3.29	2.92
Self	.32	.52	.19	.48	.50
Father	.53	.62	.50	.62	.25
Mother	.35	.52	.15	.19	.25
Siblings	5.03	4.43	4.35	3.33	3.96
Incoming Positive					
Nobody	.47	.57	.19	.33	.08
Self	.47	.57	.42	.14	.17
Father	1.21	1.24	2.08	1.86	2.58
Mother	2.32	1.48	2.73	2.33	3.08
Siblings	3.26	3.67	2.77	3.09	2.79
Incoming Negative					
Nobody	1.79	1.10	2.23	3.00	2.46
Self	.53	.67	.08	.33	.04
Father	.85	.52	.85	1.00	.67
Mother	.62	1.10	.73	.67	1.71
Siblings	3.91	4.38	1.65	2.76	2.83
Dependency					
Nobody	.56	.57	.31	.81	.42
Self	.82	.86	.88	.67	.21
Father	.94	1.19	2.08	1.67	2.17
Mother	3.18	2.38	3.50	3.67	4.46
Siblings	1.91	2.48	1.08	1.05	.83

The latter category is used to place statements that the child cannot or will not attribute to others in the family.

Tables 25 and 26 presents the mean number of references to the five members of the family falling within the five categories of feeling. Table 25 contains the data from the first administration of the instrument and Table 26 from the spring administration in Phase IV.

Data Relevant To Individual Instruments

In this section the focus will be on data related to an understanding of the particular instrument being discussed and characteristics which may be of interest to other researchers in studying preschool and kindergarten age children. The instruments will be discussed under the same framework and in the same order as in the first half of this chapter. Because of differences in the scheduling of the various instruments and the inappropriateness of certain types of analysis for some of the instruments, the data is not the same for all instruments.

Basically, means and standard deviations of the total sample for each of the four testing phases, means of age groups (based on multiples of 3 months, as of the first three months of kindergarten) for the three phases during the kindergarten year, reliability estimates, interphase correlations for each test and intercorrelations among tasks within a task category will be presented.

Category 1: Ability To Make Discriminations

Table 27 presents the total group means and standard deviations for all four phases. The approximate* mean age of this total group is 5 years, 2 months during Phase 1, 5-5 during Phase 2, 5-8 during Phase 3 and 5-11 during Phase 4. The sample group is split approximately* 60% - 40% on the dichotomies of urban-small town-rural, disadvantaged-non-disadvantaged, Negro-white and non-Head Start-Head Start.

The means, with the exception of the Phase 2 Haptic measure, demonstrate consistent growth or development over the course of the study. The relative stability of the standard deviations across the four phases, again excepting the one Haptic measure, should be noted. For those tests for which age norms are available, specifically the Columbia and the Frostig subtests, the mental or perception age corresponding to each mean is presented. It should be noted that the change between phases 1 and 4 is ten months on the Columbia, fifteen on the Figure-Ground subtest and fourteen on the Spatial Relations subtest. This compares quite favorably to the nine month spread in chronological age between these two phases.

Internal consistency reliability estimates** calculated by the Hoyt analysis of variance procedure resulted in estimates of .95 for the Columbia and .87 for the Haptic. Corresponding data for the Frostig subtests is not available. The correlations between successive phase

---

*Figures given are approximate because the several means and standard deviations are based upon different N's.

**Based on a sample of over 150 Head Start children not included in this study.

Table 27. Means and standard deviations of four tests in the Ability to Make Discriminations category for four phases of testing.

Test	Phase	N	Mean	Test Age**	Standard Deviation
Columbia	1	59	30.81	4-6	12.26
	2	150	38.51	4-11	14.43
	3	122	43.11	5-1	14.33
	4	124	47.08	5-4	13.90
Frostig Figure-Ground*	1	59	40.04 (3.98)	3-9	13.49 (4.87)
	2	148	47.23 (6.28)	4-4	14.32 (7.36)
	3	103	55.63 (9.31)	4-9	11.00 (6.29)
	4	95	57.96 (10.91)	5-0	11.66 (6.03)
Frostig Spatial Relations*	1	59	24.39 (.85)	4-7	13.78 (1.46)
	2	148	28.56 (1.48)	4-10	14.06 (1.69
	3	103	36.62 (2.77)	5-4	16.42 (2.28)
	4	95	38.17 (3.54)	5-9	15.19 (1.97)
Haptic Visual Matching	1	18	10.89		2.22
	2	148	10.19		3.84
	3	116	11.38		2.60
	4	93	12.21		2.35

*Figures in parentheses are means and standard deviations based on scoring protocols according to directions in the Frostig manual rather than by procedures described in Appendix A.

**Mental or perception age based on norms in respective test manuals.

Table 28. Mean scores on four tests in the category, Ability to Make Discriminations, for five age groups during three phases of testing.

		Age** During Phase 2				
Test	Phase	5-0	5-3	5-6	5-9	6-0
Columbia	2	33.79	35.71	40.92	43.11	41.08*
	3	39.90*	40.19	44.83	46.13	47.75*
	4	45.53	43.81	48.88	-	53.41*
Frostig Figure-Ground	2	44.56*	44.20	48.18	51.78	52.20*
	3	55.59	54.65	55.63	-	55.56*
	4	56.80	56.63	-	61.16	59.48*
Frostig Spatial Relations	2	23.68*	26.28	30.89	27.95	37.62*
	3	30.51	33.02	39.13	-	43.64*
	4	32.18	36.35	-	42.99	48.28*
Haptic Visual Matching	2	9.17	10.24	10.60	10.11	10.54*
	3	11.55	11.17	11.57	11.42	11.18*
	4	12.83	-	12.38	11.67	12.25*

*$10 \leq N < 18$; all other estimates based on $N \geq 18$.

**Note that ages increase three months between Phases so that ages during Phase 3 are three months greater than in heading and during Phase 4 they are six months greater.

administrations of each instrument may be thought of as test-retest reliability estimate. The correlations between Phases 1 and 2 and between Phases 2 and 3 are .79 and .80 for the Columbia, .74 and .80 for Figure-Ground, .71 and .84 for Spatial Relations and .71 and .35 for the Haptic. The first of these numbers would be based upon the Head Start groups only while the second is based upon all five groups. While these test-retest figures are considerably lower than the internal consistency estimates they do indicate that, on all except the Haptic, fairly stable traits were being measured.

The mean scores of age groups (according to ages during the second phase of testing) are presented in Table 28. Data for Phase 1 are not given because the sample sizes for the estimates were too small.

The relationships among the four measures is indicated by the correlations given in Table 29.

Table 29. Correlations between pairs of tests in the Ability to Make Discriminations category for Phases 2 and 4.

	Phase	Columbia	Figure-Ground	Spatial Relations
Figure-Ground	2	.57		
	4	.67		
Spatial Relations	2	.51	.56	
	4	.58	.60	
Haptic	2	.18	.43	.39
	4	.46	.51	.48

Table 30. Means and standard deviations of five tests in the Basic Comprehension Skills category for four phases of testing.

Test	Phase	N	Mean	Test Age	Standard Deviation
French PTI, Picture Vocabulary	1	59	13.81	4-11	4.76
	2	149	15.14	5-1	5.40
	3	116	17.79	5-8	4.65
	4	106	19.02	6-0	4.81
ITPA, Visual Decoding	1	60	8.17	4-5	4.20
	2	149	9.49	4-11	4.28
	3	118	11.14	5-7	3.81
	4	95	12.10	5-10	3.19
ITPA, Visual-Motor Association	1	60	8.72	3-11	4.81
	2	149	9.90	4-4	4.61
	3	118	10.93	4-8	4.33
	4	95	13.50	5-7	3.91
ITPA, Auditory Decoding	1	60	15.32	4-8	5.87
	2	149	16.87	4-11	6.70
	3	118	19.00	5-5	6.32
	4	95	20.09	5-8	6.45
Symbol Recognition, Part I	1	60	21.02		4.28
	2	150	21.95		5.20
	3	124	23.82		4.52
	4	120	24.73		4.34

Correlations among the Columbia and the Frostig subtest are moderately high, all being in the range from .51 to .68. Figures for the Haptic are somewhat lower indicating less commonality with the other three.

Category 2: Basic Comprehension Skills

Summary data for the five tests or subtests in this category are presented in Table 30. The general picture is similar to that for the preceding category with steady increments in mean scores across phases and relatively stable standard deviations as the means change.

On those tests for which norms are available mental or language ages corresponding to the group means have been estimated. The change on the Picture Vocabulary subtest is thirteen months while those for the three ITPA subtests, Visual Decoding, Visual-Motor Association and Auditory Decoding are seventeen, twenty and twelve months respectively. The reader should be cautioned against too optimistic an interpretation of these gains as the norm samples are small and the corresponding norms probably somewhat unstable. Internal consistency estimates of reliabilities and test-retest correlations are given in Table 31. The internal consistency figures range from .77 to .91.

Table 31. Reliability estimates of five tests in the Basic Comprehension Skills category.

Test	Internal Consistency	Test-Retest Phases 1 & 2	Test-Retest Phases 2 & 3
Picture Vocabulary	.81	.60	.56
Visual Decoding	.82	.52	.51
Visual-Motor Association	.85	.40	.49
Auditory Decoding	.91	.47	.48
Symbol Recognition	.77	.50	.65

All the test-retest values are between .40 and .65. These test-retest values are considerably less than those obtained for the tests in category 1, probably reflecting the greater impact of kindergarten on comprehension skills.

Correlations among the five tests in this category are given in Table 32. These correlations are modest indicating that the subtests, to the extent that they measure comprehension skills, are measuring different facets of these skills.

Table 32. Correlations between pairs of tests in the Basic Comprehension Skills category for Phases 2 and 4.

	Phase	Picture Vocabulary	Visual Decoding	Visual-Motor Association	Auditory Decoding
Visual Decoding	2	.51			
	4	.57			
Visual-Motor Association	2	.31	.42		
	4	.23	.18		
Auditory Decoding	2	.52	.40	.48	
	4	.33	.36	.24	
Symbol Recognition	2	.43	.29	.31	.35
	4	.54	.57	.23	.33

The mean scores for subjects grouped by age at time of testing appear in Table 33. It should be noted that the diagonal rows moving down and to the left are means for groups of equal age but differing by three month increments in school experience. Looked at from this perspective one notes a general, but irregular, pattern of increase on these tasks due to school experience as well as increase in age.

Table 33. Mean scores on five tests in the category, Basic Comprehension Skills, for five age groups during three phases of testing.

Test	Phase	Age at Phase 2 Testing				
		5-0	5-3	5-6	5-9	6-0
French PTI, Picture Vocabulary	2	14.25	13.78	15.08	18.0	17.54*
	3	19.10	17.28	17.22	17.38	19.63*
	4	-	17.60	19.65	20.68	-
ITPA, Visual Decoding	2	9.88	8.11	9.94	10.47	10.62*
	3	11.86	10.14	12.29	11.21	12.00*
	4	13.05	-	11.90	12.47	12.00*
ITPA, Visual-Motor Association	2	17.33	8.04	10.47	11.53	9.38*
	3	11.86	10.30	12.19	11.32	9.27*
	4	14.00	-	13.70	12.16	13.42*
ITPA, Auditory Decoding	2	17.21	16.67	16.15	20.21	14.61*
	3	16.95	18.23	19.19	20.21	23.82*
	4	18.32	-	19.53	20.63	22.75*
Symbol Recognition, Part I	2	21.58	21.22	22.20	22.10	23.92*
	3	22.80*	23.12	24.50	23.79	-
	4	24.90	24.19	24.70	-	26.80*

*$10 \leq N < 18$; all other estimates based on $N \geq 18$.

Category 3: Basic Quantitative Skills

In this category, only the Size and Number subtest of the French PTI was administered in each of the four phases. Tables 34 and 35 summarize the mean and standard deviation data for this instrument for total and age group samples. The means increase steadily across phases with the total increase corresponding to a mental age change (estimated from norm tables in manual) of approximately thirteen months.

Table 34. Means and standard deviations on the French PTI, Size and Number subtest in the Basic Quantitative Skills category for four phases of testing.

Test	Phase	N	Mean	Test Age	Standard Deviation
French PTI, Size and Number	1	59	7.98	4-3	3.91
	2	149	9.15	4-7	4.55
	3	116	11.78	5-1	5.57
	4	106	13.08	5-4	5.81

Table 35. Mean scores on the French PTI, Size and Number subtest for five age groups during four phases of testing.

		Age at Phase 2 Testing				
Test	Phase	5-0	5-3	5-6	5-9	6-0
French PTI Size and Number	2	8.00	8.42	9.35	9.95	11.77*
	3	9.05	12.50	11.35	13.47	14.18*
	4	-	11.79	13.42	15.21	-

*$10 \leq N < 18$; all other estimates based on $N \geq 18$.

The data in Table 35 indicates consistent increases in mean scores corresponding to both higher age and added school experience.

The internal consistency reliability estimate for Size and Number is .81. Test-retest correlation between Phases 1 and 2 was .62 and between Phases 2 and 3 was .81.

Category 4: General Ability and Knowledge

In this category only the General Information Test was administered more than once. As seen by the data in Tables 36 and 37 there is, again, a steady increase in means of the total group across phases with relatively stable estimates of the standard deviation. However, in Table 37 there is somewhat less variability across age levels within any given phase. This is probably due to the extreme school relatedness of the tasks on this test. That is, the tasks or bits of information called for are ones that the kindergarten program systematically attempts to teach or expose the child.

The test-retest correlation on this test was .78 between Phases 1 and 2 and .82 between Phases 2 and 3. Because of the diverse nature of information asked for on this test an internal consistency estimate was not appropriate.

No additional data is available on two of the other three tests in this category, the Caldwell Preschool Inventory and the Pintner-Cunningham.

The final test in this category was the Wechsler Pre-School and Primary Scale of Intelligence. This instrument was subjected to analyses beyond that of the other instruments because (1) it is the

Table 36. Means and standard deviations on the General Information Test for four phases of testing.

	Phase	N	Mean	Standard Deviation
General Information	1	59	15.97	6.21
	2	149	18.41	6.92
	3	119	21.35	7.22
	4	102	23.72	6.81

Table 37. Means of five age groups on the General Information Test for three phases of testing.

		Age at Phase 2 Testing				
	Phase	5-0	5-3	5-6	5-9	6-0
General Information	2	16.96	17.00	18.50	21.05	21.77
	3	18.91	22.95	20.91	22.74	22.46
	4	-	22.32	24.54	24.39	-

single most complete measure of cognitive functioning used in the study and (2) its potential popularity among school psychologists and others responsible for individual diagnosis of learning problems demands more information.

In Appendix B the item by item mean scores by group are listed. This information has been placed in an Appendix because of the bulk of the tables needed to present the data. This data was studied to (1) determine the progression of item difficulty in each subtest and (2) to identify any items that present the Negro urban-disadvantaged group with undue difficulty. The procedure yielded a mean value for each item. This allowed a comparison of the item difficulty between the disadvantaged and non-disadvantaged groups as well as the rural and urban groups. As expected, the items generally proceeded from "easy" to "difficult" on the subtests. Occasionally an item proved to be more difficult than those subsequent to it. The Information subtest presented several examples of greater-than-usual difficulty for Groups A and B: (Item 5) "What lives in water?" and (Item 9) "What shines in the sky at night?" Item 4 of the Arithmetic subtest also presented greater difficulty for these children than either Items 5 or 6. The subject is presented a picture showing four bowls of cherries. He is then asked "Which bowls have the same number of cherries?" A lack of understanding of the word "same" is no doubt a major problem for these children. A last example resides with the Comprehension subtest. Item 2: "Why do you need to wash your face and hands?" presents greater difficulty for these subjects than either Items 3, 4 or 6. Other less apparent examples are present but will not be discussed. For the most part, however, the sequence of item difficulty is similar to the WPPSI subtest sequences.

The data presented earlier in this chapter indicated that Groups A and B were functioning at a level considerably below the other groups on all of the WPPSI subtests. To investigate the possibility that test characteristics might be different for these black urban disadvantaged groups the data was pooled and intercorrelations among subtests calculated. This data is presented in Table 38. Those correlations marked with a single asterisk signify a deviation of -.15 or more from the correlation presented in the WPPSI manual (Table 15, p.30). The double asterisked correlation signifies a deviation of +.15 from

Table 38. Intercorrelations among the ten subtests of the WPPSI. (Based on 54 children in the urban, disadvantaged Negro groups.)

I.	V.	A.	S.	C.	AN.H.	P.C.	M.	GEO.D.	B.D.	Subtest
	.407	.666	.502	.572	.431	.455	.392	.571**	.252*	I.
		.269*	.289*	.469	.230	.320*	.201*	.292	.242*	V.
			.451	.463	.424	.382	.469	.358	.266*	A.
				.482	.414	.317	.389	.216*	.231*	S.
					.352	.359	.365	.310	.188*	C.
						.470	.330	.339*	.366	AN.H.
							.281*	.351*	.373*	P.C.
								.283*	.438	M.
									.350*	GEO.D.
										B.D.

I = Information
V = Vocabulary
A = Arithmetic
S = Similarities
C = Comprehension
AN.H. = Animal House
P.C. = Picture Completion
M = Mazes
GEO.D. = Geometric Design
B.D. = Block Design

*Correlation deviates -.15 or greater from value presented in WPPSI Manual (Table 15, p. 30).

**Correlation deviates +.15 or greater from value presented in WPPSI Manual (Table 15, p. 30).

the correlation presented in the manual. In general, the inter-test correlations for this urban-disadvantaged group were of a smaller magnitude than those presented in the manual. The reason for this is not entirely clear but may be due to the homogeneity of the sample. If one identifies all correlations which differ from those in the WPPSI manual by some arbitrarily large amount, say ±.15 which is at least one standard error, then several interesting patterns emerge from this overall tendency toward lower correlations. Eighteen of the correlations deviate from Wechsler's by ±.15, 17 in the negative direction and one in the positive direction. Of these 17 correlations, 16 involve the Vocabulary, Geometric Design and Block Design subtests. One plausible explanation for the decrement in correlations involving the Vocabulary subtest is as follows: of all the WPPSI subtests, the Vocabulary mean and standard deviation is the lowest which in turn serves to depress the magnitude of the correlations. The Block Design subtest uniformally correlates with every verbal subtest to a lesser degree when compared to the manual values. That the Block Design subtest is probably the best nonverbal estimate of abstracting ability has been recognized in the research involving the Wechsler Intelligence Scale for Children (WISC). Likewise the verbal subtests appear to involve a moderate degree of verbal abstraction. This common "factor" of abstraction may account for a sizable proportion of the Block Design-Verbal subtests relationship on the WISC. The designs presented early in this WISC subtest most likely require a higher degree of abstraction than the initial 4-5 designs of the WPPSI. The majority of our sample of urban-disadvantaged children successfully completed only the first five designs. These designs may be considered near the concrete end of an "abstract-concrete" continuum of design tasks. Therefore, by virtue of completing the less abstract of the designs, the correlation of performance with the more abstract verbal subtests is reduced. The final aspect to notice in Table 1 is the lowered correlations of the Geometric Design subtest with its performance counterparts. The reason for this decrement is not at all clear but should create some interesting speculations.

Table 39 compares the mean scale scores of the pooled urban, Negro, disadvantaged group (A plus B) with the small town-rural, white, disadvantaged group (D). Standard errors and interval estimates of the means are also presented.

Lastly this table presents the mean Verbal, Performance and Full Scale I.Q.'s of the three groups. It is readily apparent that the urban-disadvantaged subjects demonstrate a marked decrement in all subtests of the WPPSI when compared to the norms. For the urban-disadvantaged group, with the exception of the Vocabulary subtest, scaled scores for the verbal subtests present a pattern of similarity. The performance scores for this group, however, present less of a distinct pattern. Their lowest subtest performance involves the vocabulary skills; the group score falling approximately 1 2/3 standard error of the mean below the other verbal subtests. It is thought that this subtest measures "one's potentiality for dealing with symbols at both the concrete and abstract level." The test appears to be "loaded with factors of auditory comprehension and is thought of as a test of verbal concept formation" (Wentland and Geeding, 1967). A decrement

Table 39. Mean scale scores on all subtests of WPPSI and mean Verbal, Performance and Full Scale IQ scores for the two disadvantaged groups.

	Groups A & B (N = 54)			Group D (N = 26)		
Subtest*	Mean	$S_{\bar{x}}$	95% CI	Mean	$S_{\bar{x}}$	95% CI
I.	7.5	.35	6.8-8.2	9.8	.55	8.7-11.0
V.	6.7	.31	6.1-7.4	8.5	.50	7.5- 9.5
A.	7.5	.39	6.7-8.2	9.9	.65	8.6-11.2
S.	7.2	.40	6.4-8.0	8.6	.55	7.5- 9.7
C.	7.2	.34	6.5-7.9	9.0	.60	7.8-10.3
AN.H.	7.3	.37	6.5-8.0	9.8	.55	8.6-10.9
P.C.	8.2	.37	7.4-8.9	9.4	.53	8.3-10.5
M.	6.8	.28	6.3-7.4	9.3	.55	8.1-10.5
GEO.D.	7.9	.50	6.9-8.9	9.62	.95	7.7-11.6
B.D.	8.2	.43	7.3-9.0	9.9	.53	8.8-11.0
Verb. I.Q.	83.2			95.0		
Perf. I.Q.	85.0			97.7		
F.S. I.Q.	82.6			95.9		

*See Table 38 for abbreviations.

in performance is also shown on the Maze subtest. To the extent that the WPPSI Maze test is similar to that of Porteus (Porteus, 1950), the "Task requires the subject to gain an overview of the whole problem and to organize and plan his procedure with a minimum waste of time and a minimum amount of path retracing. . .The test requires a fair degree of perceptual orientation and analysis")Horrocks, 1964).

In contrast to the performance scores of the pooled Groups A and B, note the tendency for greater clustering of the performance scores for Group D. With the two exceptions (Vocabulary and Similarities) that require a greater degree of abstract verbal ability, this latter group performs in a manner similar to Wechsler's norm group.

Relative to the Verbal, Performance and Full Scale I.Q.'s, a 12-13 point difference between the two disadvantaged groups is in evidence. The small town-rural disadvantaged subjects lie within the average range of functioning while the urban disadvantaged demonstrate a significant decrement in all areas measured by the WPPSI. No doubt a large proportion of the difference relates to differences in early stimulation and range of experiences.

Category 5: Cognitive Style

Of the three measures originally intended for use in the Cognitive Style category only the Block Sorting Task seemed to provide valuable information. Both the Matching Familiar Figures Test and the Haptic Visual Matching Test, in the modifications used here, provided questionable results. Table 40 presents the mean response time data for the two instruments. Note that the means, particularly on Matching

Familiar Figures are very constant across phases while the standard deviations vary considerably, particularly for Phase 2 where they are two and three times as large as for the other phases.

Table 40. Mean response times on the Haptic Visual Matching Test and the Matching Familiar Figures Test in the Cognitive Style category during four phases of testing.

	Phase	N	Mean Time	Standard Deviation
Haptic Time (in seconds)	1	18	96.2	22.7
	2	148	103.5	87.3
	3	116	86.2	26.9
	4	93	98.1	36.0
Matching Familiar Figures Time (in seconds)	1	18	68.8	25.3
	2	151	69.8	54.4
	3	133	62.5	26.4
	4	99	67.8	24.1

The third task, the Block Sorting Task, was only administered once and there is no additional data to that presented in the early discussion of group differences.

Category 6: Affective Traits

Means and standard deviations on the Delay of Gratification Task and the Illinois Index of Self-Derrogation are presented in Table 41. The general tendency on the Delay of Gratification measure is in the direction of more delay or increased willingness to wait for the larger reward. The authors have some doubts about the meaningfulness of this trend and suspect that it is highly specific to one and five cent Tootsie Rolls. However, the Phase 2 figure of

Table 41. Means and standard deviations on two tasks in the Affective Traits category for four phases of testing.

Task	Phase	N	Mean	Standard Deviation
Delay of Gratification	1	53	.19	.40
	2	148	.29	.46
	3	127	.39	.49
Illinois Index of Self-Derrogation	1	14	17.79	4.63
	2	146	20.76	4.60
	3	73	20.41	5.04
	4	98	22.55	4.82

Table 42. Mean proportions or scores on two techniques in the Affective Trait category for five age groups.

		Age at Phase 2 testing				
Instrument	Phase	5-0	5-3	5-6	5-9	6-0
Delay of Gratification	2	.38	.30	.29	.21	.23
	3	.44*	.47	.38	.32	.27*
Illinois Index of Self-Derrogation	2	19.75	20.44	21.04	20.58	22.92
	3	20.70*	21.65	19.30	19.17	-
	4	23.28	22.61	22.00	-	-

* 10 < N < 18

29 percent choosing to wait for the larger reward and 71 percent choosing the smaller, but immediate, reward should be a good estimate of how kindergarten children will divide on this type of task. In Table 42 the results indicate a tendency for younger children to wait and older children to request the immediate reward.

On the self-concept or self-perception measure, the Illinois Index of Self-Derrogation, there is a consistent trend toward a more positive concept as a result of the kindergarten experience. We attribute the change to experience rather than age because the means of the age groups, as presented in Table 42, indicate no consistent relationship between age and self-perception.

Category 7: Family Relationships

The Bene-Anthony Family Relations Test used as the measure in this category does not provide any single summary score. Rather, it results in the type of twenty-five socre profile presented in Table 25 and 26. Because of the large number of summary numbers in those two tables (actually 250 mean frequencies based on fewer than 150 individuals) one begins to have degrees of freedom problems that could easily lead to a number of interesting but probably spurious relationships. Rather than over-analyze we have chosen to restrict the data on this measure to that in Tables 25 and 26.

Summary

The essence of this study is description, description of diverse groups over time on a variety of instruments which aim at measuring a smaller subtest of cognitive and related affective traits. The result has been a veritable mountain of data which has been presented in this chapter through forty-one tables and twenty graphs. Where possible an attempt has been made to indicate the precision or stability of the mean estimates by reference to a measure of error.

While the next chapter will attempt to discuss what these writers see as some of the more interesting aspects of the data, we believe that the major value of the report is in this chapter. Interpretation or extraction of meaning from a set of data is in part a function of the problem with which one approaches the data. We encourage the reader to reexamine the many tables and graphs in this chapter with this problem in mind.

CHAPTER V

INTERPRETATION OF RESULTS

In Chapter IV the results of the study were presented in considerable detail, first, as they described the five groups being studied and second, as normative data about the instruments and technique that should be of value to prospective users of them with children of kindergarten age. This chapter will attempt to synthesize the information in Chapter IV, elaborating on the distinctive characteristics of the five groups of children and emphasizing the similarities and differences among them.

Group characteristics and relative strengths and weaknesses are difficult to obtain from a study of the data as presented in Chapter IV. The various tests are scored on different scales, and the extreme amount of information pertaining to changes over time and for children of different ages can only tend to confuse the reader. For purposes of emphasizing distinctive features of the various groups' performances, the following discussion will focus on the results of the final or spring phase of testing. This is appropriate because the spring data is a description of these groups and their abilities as they are about to enter the more formal learning program of first grade. Differences that existed during earlier phases but subsequently have disappeared are not as important as those which still exist at this first crucial transition point in the child's education.

To allow for comparisons across instruments, the group means have been standardized and are presented in Tables 43 and 44 and in Figures

21 through 25. The standardization procedure consisted of subtracting the grand mean (actually a mean of the five group means) from the respective group means and dividing the result by the standard error claculated for the test. These standardized means are then presented in the seven category organization used in the last chapter. While the exact values are presented in the tables, our discussion will focus upon the five graphs of subsets of the measures. Lines have been drawn connecting the points for each group, not to indicate trends as in the graphs in Chapter IV, but to make it easier to see the response pattern for each of the five groups.

Before interpreting the results category by category, some general remarks are in order. The two groups which are the major focus of the study are Groups A and D, the two Head Start groups. The reader is reminded that Group B was chosen as a comparison or control group for Group A, and that it consisted of children living in the same neighborhoods as the Group A children, attending the same kindergarten classes but, while eligible for summer Head Start, not participating in that program. No comparable control group could be found for the small town-rural Head Start group. Groups C and E were chosen as control groups of a different type. They represented the types of middle class or non-disadvantaged children that public schools seem best equipped to handle. As such, they were seen as providing base line data against which to measure the relative educational disadvantaged of the poverty groups A, B and D.

Studying the overall performances of the five groups leads to the following generalizations. Group E, the white, middle-class, small town-rural group, is consistently the top performing group. Adopting the convention that a difference of less than two standard errors is not worthy of discussion, there is no measure of maximal performance on which E is not at the top. Similarly, there is no measure on which A and B are not the bottom two or among the bottom group. This is not to say that the profiles for these three groups are, or should be, horizontal lines because the variability between groups differs considerably from test to test. Furthermore, there is a consistent tendency for Group A to outperform Group B. This difference, however, does not appear to be due to the participation in Head Start by Group A. Examination of the difference in performances between Group A, tested during Phase 1, and Group B, tested during Phase 2, leads to the conclusion that Group B is a less able group and that the most disadvantaged children are the very ones that voluntary programs such as Head Start are likely to miss.

The other two groups, C and D, have profiles that are far less consistent. On certain measures they perform at an approximately equal level somewhere between that of Group E and Groups A and B. On other measures, one or both may equal either the relatively high performance of Group E or the lower performance of Groups A and B. Group D certainly does not display the cognitive deficits that distinguish the urban disadvantaged children. In fact, these small town-rural Head Start children demonstrate no greater overall cognitive or educational disadvantage or deficit than do the children in the black, urban non-economically disadvantaged group.

Figure 21 summarizes the Phase 4 performance of the five groups

Table 43. Standardized mean scores for five groups of children on eleven instruments during final phase of testing.

Category/Test	Mean of Means	Standard Error	A	B	C	D	E
Discrimination abilities							
Columbia	47.18	2.21	-2.88	-5.92	2.18	2.36	4.26
Figure-Ground	57.54	2.39	-1.40	-3.94	3.08	.78	1.03
Spatial Relations	37.96	2.80	-2.95	-2.13	-1.48	2.01	4.55
Haptic	12.20	.65	- .50	-1.02	.74	0.00	.78
Comprehension skills							
Picture Vocabulary	18.98	.86	-1.70	-6.41	3.08	2.05	2.98
Visual Decoding	12.02	.75	-1.09	-3.42	2.56	1.17	.78
Visual-Motor Association	13.48	.89	.16	- .11	1.01	-1.84	.78
Auditory Decoding	19.93	1.17	-3.49	-2.64	1.39	- .08	4.82
Symbol Recognition I	24.80	.84	-1.60	-4.30	1.52	1.49	2.89
Symbol Recognition II	29.38	.46	-1.36	-2.70	1.78	1.24	1.04
Quantitative skills							
Size and Number	13.01	.86	-3.18	-5.95	1.02	2.80	5.31
Conservation of: (Elkind)							
Number	.45	.11	.36	- .82	- .18	- .55	2.00
Volume (2 parts)	.26	.10	.60	.30	- .20	- .90	.30
Volume (shape)	.27	.10	- .90	.20	1.10	-1.50	1.10

Table 44. Standardized mean scores for five groups of children on seven instruments during final phase of testing.

Category/Test	Mean of Means	Standard Error	A	B	C	D	E
General ability							
General Information	23.87	1.13	-3.57	-5.94	2.07	2.93	4.49
Pintner-Cunningham	92.78	2.89	-2.36	-1.58	- .30	.70	3.54
Caldwell	60.34	1.87	-4.72	-7.99	2.81	3.56	6.34
WPPSI							
Verbal	43.35	2.28	-2.41	-3.92	1.78	1.22	3.33
Performance	45.25	2.38	-2.28	-3.22	1.19	1.02	3.29
Cognitive style							
Hatpic Time	97.20	14.10	.19	- .62	- .04	.24	.23
Affective traits							
Illinois Index of Self-Derrogation	23.03	1.05	-2.44	-2.43	1.76	.83	2.28
Delay of Gratification	.38	.09	- .38	- .04	3.18	-2.16	- .06

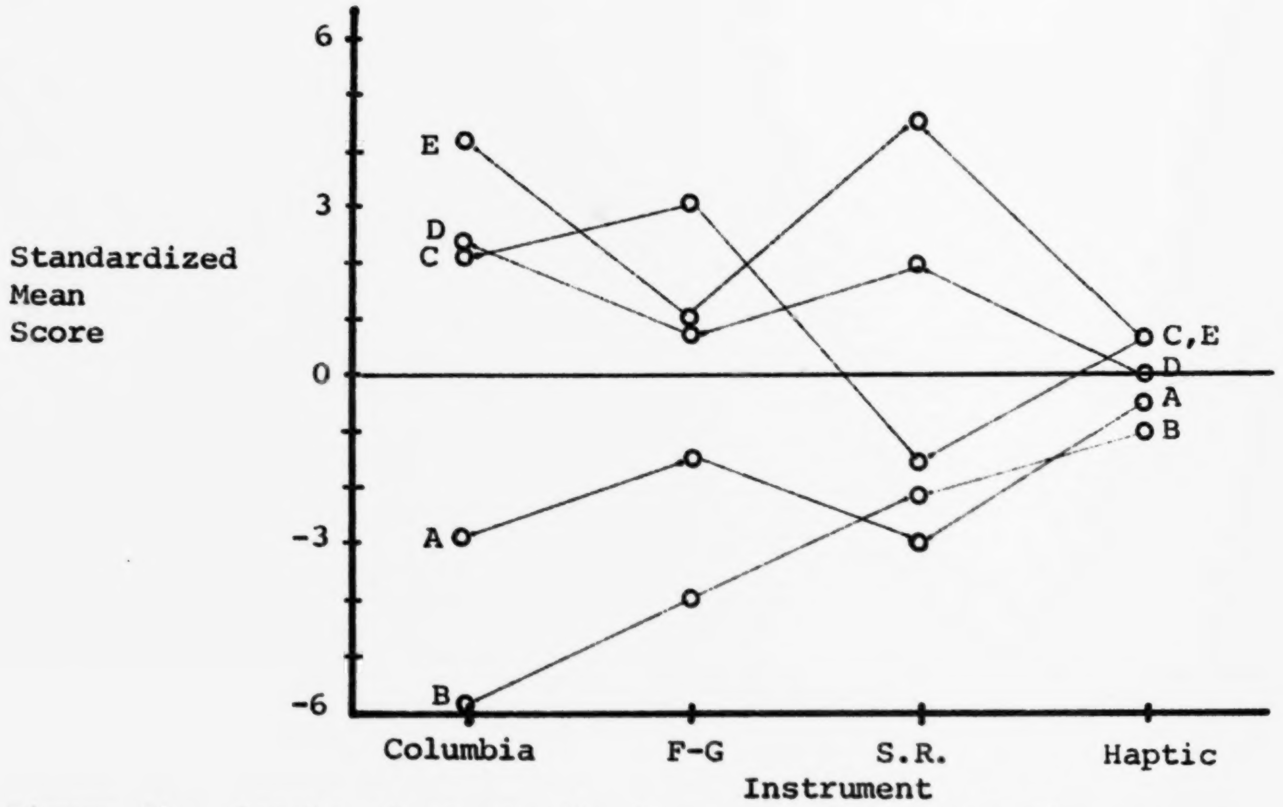

Figure 21. Graphs of standardized mean scores of the five groups of children on four tests in the Ability to Make Discriminations category. (During Phase 4 of testing.)

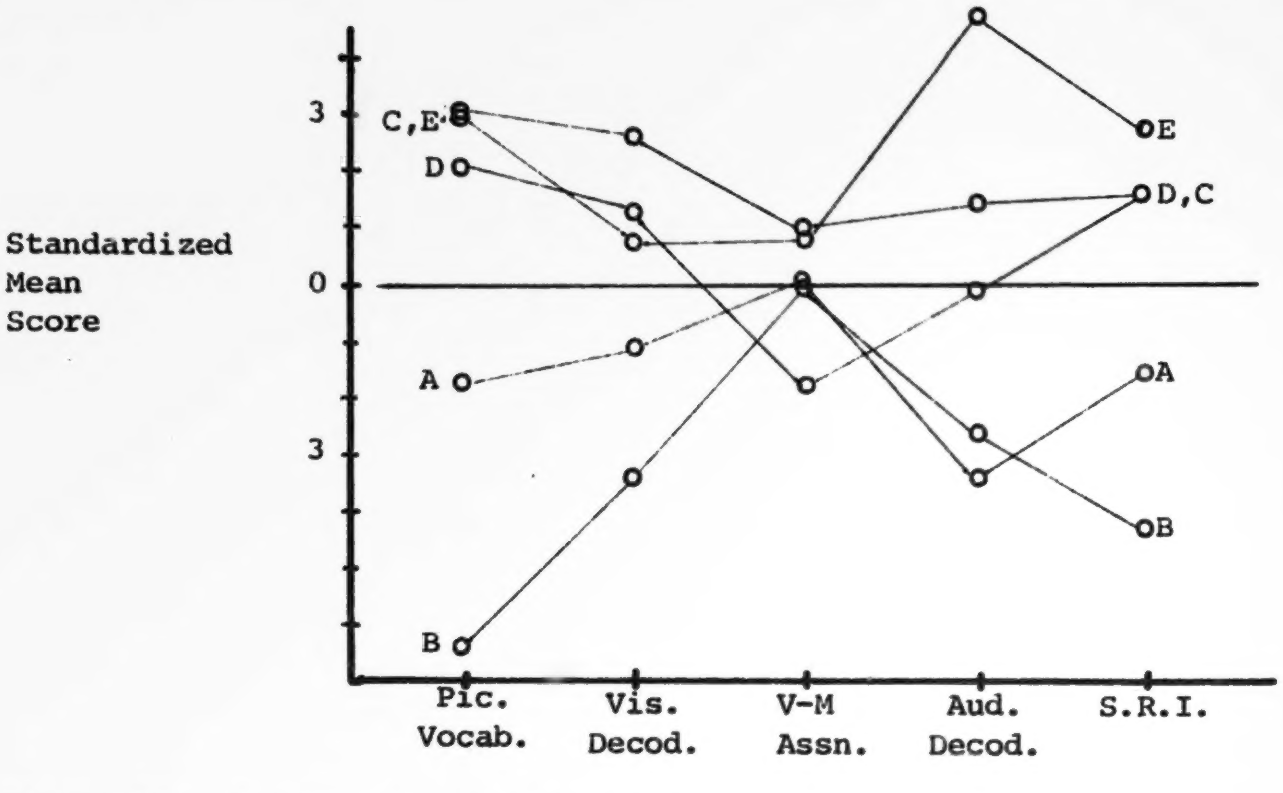

Figure 22. Graphs of standardized mean scores of the five groups of children on the five tests in the Basic Comprehension Skills category. (During Phase 4 of testing.)

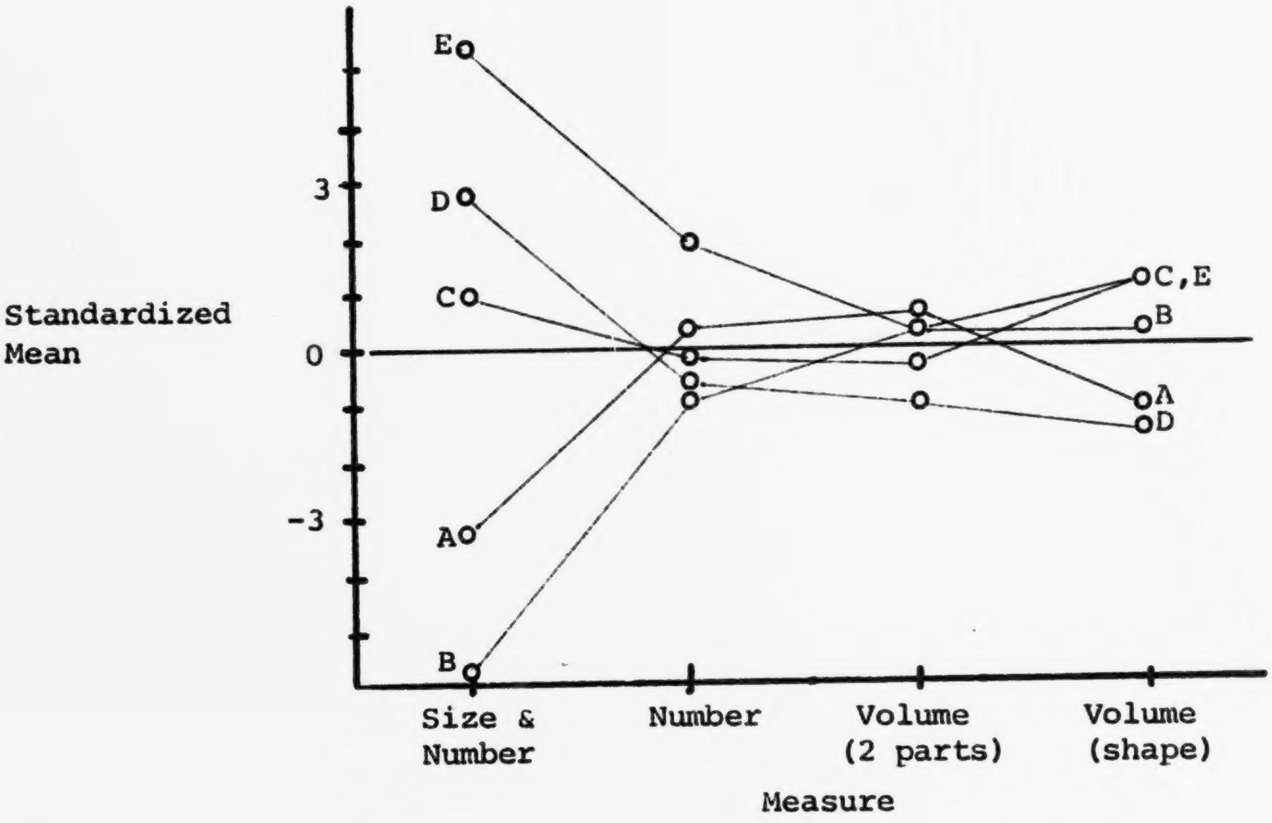

Figure 23. Graphs of standardized mean scores and proportions for the five groups of children on the two tests in the Quantitative Skills category. (During final phase of testing)

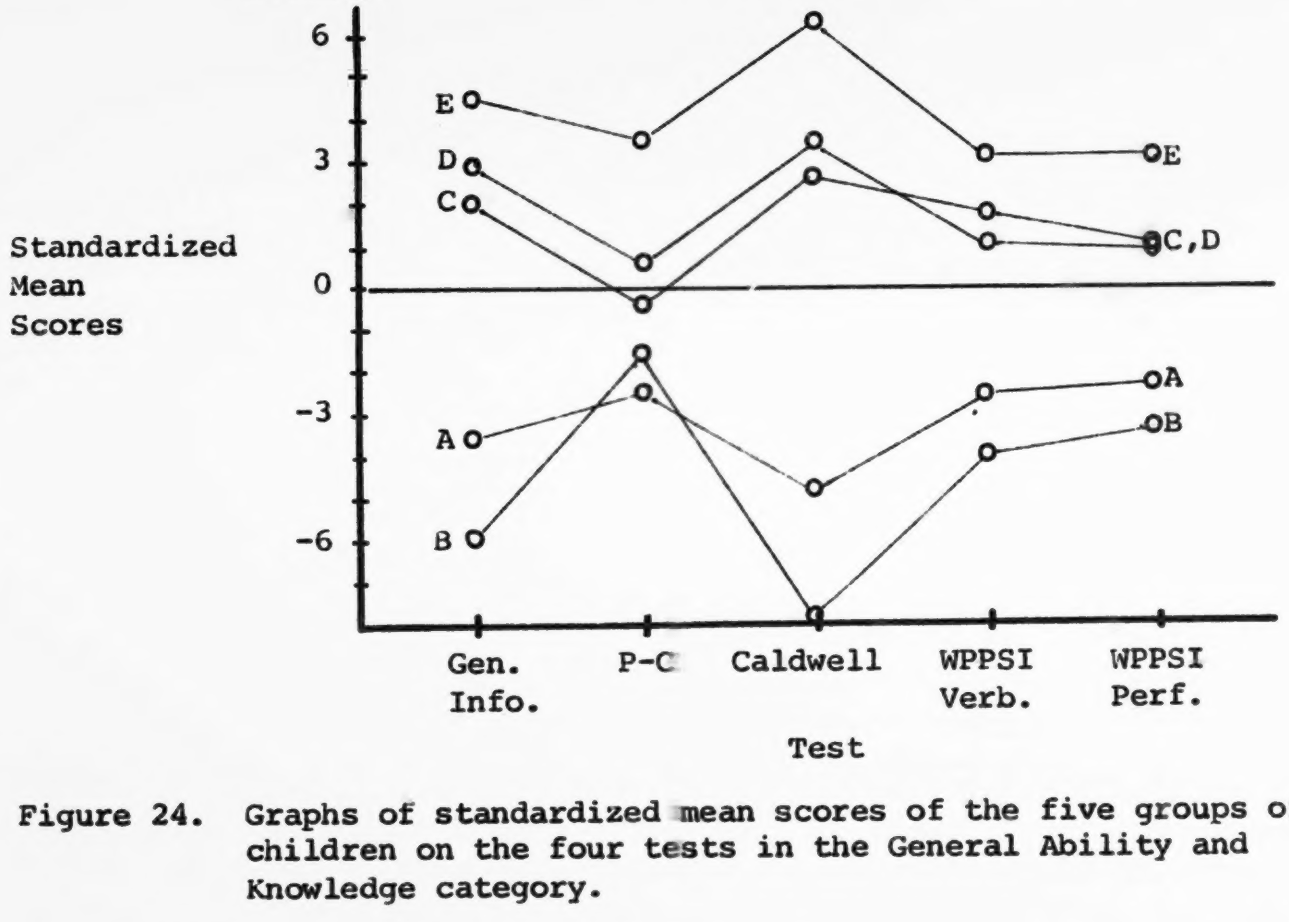

Figure 24. Graphs of standardized mean scores of the five groups of children on the four tests in the General Ability and Knowledge category.

*In standard score units from WPPSI manual.

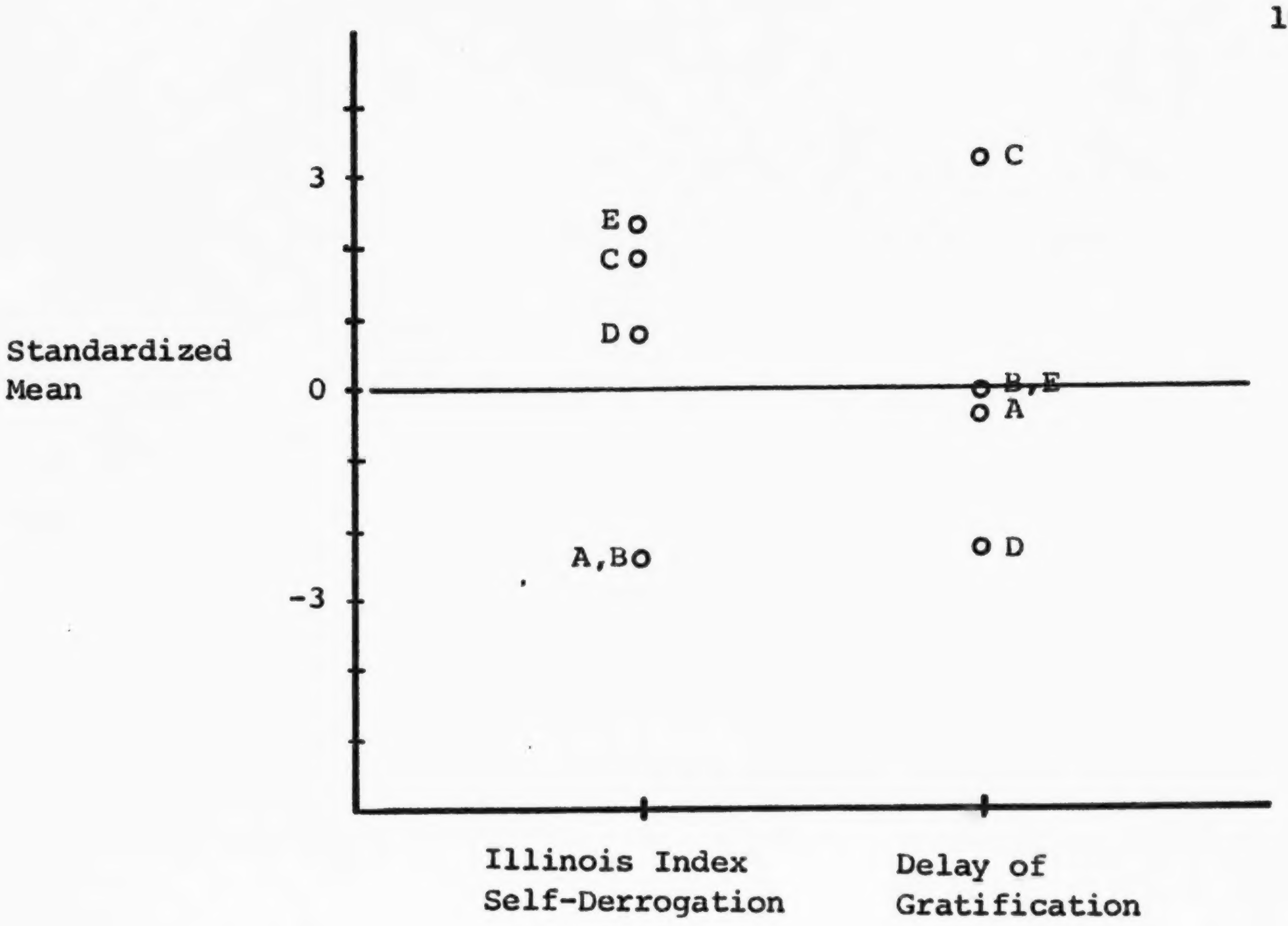

Figure 25. Graphs of standardized means and proportions for the five groups of children on the Affective Traits category tests. (During final phase of testing)

on the tasks in the Ability to Make Discriminations category. The instruments included in this category were: (1) Columbia Mental Maturity Scale, (2) Frostig Figure-Ground, (3) Frostig Spatial Relations, and (4) Haptic Visual Matching Test. Note the separation between Groups A and B on the one hand and Groups D and E on the other. Only on the Haptic measure is there no measurable difference. Group C clearly is performing on a level comparable to Groups D and E on all except the Spatial Relations subtest on which its mean is at the general level of Groups A and B.

It is interesting that the profiles for the two small town-rural groups have generally the same shape with relatively strong performances on the more global Columbia and on the Spatial Relations subtest, while the three urban groups also tend to have similar profiles, but with relatively stronger performance on the Figure-Ground and Haptic test. This similarity in shape is particularly pronounced between Groups A and C. But while the differences between Groups D and E are, at most, 2.54 standard errors, the differences between A and C are 5.08 standard errors on the Columbia and 4.48 on the Frostig Figure-Ground. Thus, on those discrimination skills measured by these two instruments, the urban disadvantaged groups manifest far greater deficits than the small town-rural disadvantaged children.

The poor showing of all three urban groups on the Spatial Relation tests would seem to call for some special attention to this area in Head Start and kindergarten curricula.

Group A shows a more-or-less even development in the discrimination and visual-perceptual skills as tapped by the selected instruments. The one exception being the stronger performance on the Haptic test.

Another approach to interpreting the results of Chapter IV is to examine the increase in group means on the standardized instruments in the battery by comparisons to the age norms provided in the several manuals. Since the norm tables only provide age equivalents for integer-valued scores, it is necessary to interpolate to find values corresponding to the non-interger means. This problem combined with difficulties with the size and representativeness of the norming samples of some of the instruments requires extreme caution in interpretation. It does, however, provide the only means of tying our results into a picture of the larger world to which we might hope to generalize.

Looking at Group A's growth in test-age (months) for only the Columbia, one sees an increase of 10 months over the 9 month evaluation period. Generally speaking, this group advanced at an average rate in the development of the relational and visual-perceptual abilities as measured by the Columbia. Even with these gains, however, Group A fell approximately 12 months below the norms at the final evaluation point.

Group B, on the other hand, demonstrated a more uneven profile with the Columbia performance falling well below the other three areas in this category. As with the previous group, the Haptic Test performance is the strongest.

This group showed a modest amount of growth (4 months) over the 6 month evaluation period on the Columbia. At the final evaluation point, Group B remained about 17 months below the norms on this test. Among other possible explanations, this suggests that the relational abilities measured by the Columbia were emphasized to a small degree in this group's day-to-day experiences.

The profile for Group C shows some interesting peaks and valleys. The performance on the Frostig Spatial Relations falls several standard score values below the other areas suggesting this to be a less well developed skill. It is in this area that the C group clusters with the two urban-disadvantaged groups (A and B). The other three discrimination skills show a development of a fairly even nature.

Over the six months of the evaluation for Group C, they demonstrated a growth on the Columbia of eight months. At the final evaluation point, this group surpassed Groups A and B by 9 months and

14 months respectively. Group C's final performance was lower than the expected performance for this group's mean age by about three months.

Although showing a relatively flat profile compared to other groups, Group D's greatest strengths are demonstrated in the abilities tapped by the Columbia and the Frostig Spatial Relations.

The profile for Group D is strikingly similar to that of Group E suggesting a similar set of experiences in the discrimination skills for the two groups.

Over the nine months of evaluation for Group D, a gain of 12 months in test age was acquired. At the final evaluation point, this group had the same performance level as did Group C (5 years 9 months), and fell about six months below age expectations as provided by Columbia norms.

Lastly, the profile for Group E shows their superiority over the other four groups in all but the Frostig Figure-Ground subtest. This group's greatest strength is shown in the Frostig Spatial Relations Test.

Group E's final performance on the Columbia was two months below the age expectations per the norms provided. The gian over the six months of evaluation was seven months indicating average development of the relational and visual-perceptual skills.

Figure 22 presents the standardized means of the five tests in the Basic Comprehension Skills category. The instruments included in this category were: (1) Picture Vocabulary - Pictorial Test of

Intelligence, (2) Visual Decoding - Illinois Test of Psycholinguistic Abilities (ITPA), (3) Visual-Motor Association - ITPA, (4) Auditory Decoding - ITPA , and (5) Symbol Recognition (abstract and concrete). The reader is again referred to Chapter IV for a detailed rationale of this categorization process. Again, the overall showing of Groups A and B is considerably poorer than that of the other three groups (with the exception of the ITPA Visual-Motor Association subtest) with Broup B consistently poorer than Group A. Groups C, D, and E cluster on all instruments except the Visual-Motor Association subtest on which Group D is 2.62 standard errors below Group E and on the Auditory Decoding subtest on which Group E is 3.43 standard errors above the mean of Group C.

The two instruments on which there is the greatest spread of performance are on the Picture Vocabulary subtest which purportedly measures listening vocabulary, and on the Auditory Decoding subtest which is essentially a listening type task requiring decoding of entire sentences.

In looking at Group A's profile, there is a more-or-less even development of the comprehension skills except that of Auditory Decoding. (This test purports to measure the ability to comprehend the spoken word.) Group A not only was the lowest of all groups on this ability, but this area was two or more standard score values below the other comprehension areas. Its greatest strength was demonstrated in Visual-Motor association (ability to relate meaningful visual symbols).

The gain of 25 months in Visual-Motor Association over the 9 month evaluation period suggests that curricular emphasis may have been placed in this area for Group A. Their performance at the final evaluation point shows Group A to be only 3 months below the "average" level when compared to the norms for this subtest. Likewise, the gain of 21 months for this group placed them only three months below the expected level for chronological age on the Picture Vocabulary task. On the Visual Decoding subtest (ability to comprehend pictures), Group A gained 17 months, but fell about 6 months below norm expectations. Auditory Decoding, as was indicated earlier, was the most difficult for this group. A gain of 6 months over the 9 month evaluation period was made. At the final evaluation point, Group A fell 15 months below the norms on this decoding task.

Group B shows a more dramatically uneven profile than the other groups. Picture Vocabulary skill falls about 3 standard score values below the profile mean. This test purports to measure language development by assessing verbal comprehension (of single words). As with Group A, their greatest single strength was the ability to relate meaningful visual symbols (Visual-Motor Association).

Although the gain on the Picture Vocabulary test was substantial (15 months) for Group B, their performance at the final evaluation point fell 21 months below an "average" performance as indicated by the test norms. This gain was 6 months less than those obtained by the remaining four groups (21 months gain). Visual Decoding likewise fell at a point 15 months below norm expectations at the final evaluation point. The gain on this subtest was 8 months over the 6 month evaluation period.

Group B made its greatest gain on the Visual-Motor Association subtest (21 months). Its performance at the final evaluation point was only 7 months below expectations for an "average" child of comparable chronological age. All the groups clustered together on this subtest - there appeared to be no meaningful difference between Group B and the non-disadvantaged groups on this task. The possibility is suggested that the curricular experiences for both Groups A and B may have included a singificant exposure to tasks emphasizing the ability to relate meaningful visual symbols.

In general, Group C enjoyed a slight superiority over all the remaining groups in the category of comprehension skills. Their profile appears evenly developed and surpasses the small town-rural non-disadvantaged Group (E) on three of the five instruments. If any strengths per se can be deduced from this profile, they would be in the Picture Vocabulary and Visual Decoding areas. . .there is no apparent weak area for Group C in this category of abilities.

In terms of gain over the 6 month evaluation period, Group C advanced 21 months on the Picture Vocabulary subtest. At the final evaluation point, their performance was 18 months above the level expected oa an "average" child of comparable chronological age. In Visual Decoding, this group gained 10 months and placed about 2 months below norm expectations

The gain on the Visual-Motor Association subtest was 13 months and at the final performance point Group C was only 3 months below "average" when compared with the test norms. In Auditory Decoding, this same group made gains of 9 months and exceeded the expected test norms for children of comparable chronological age by 2 months. No test ages are provided for the Symbol Recognition test, but a glance at the profile shows no difficulty on these tasks for Group C.

The profile for Group D shows a moderate amount of unevenness with the Picture Vocabulary skills showing the greatest strength by a small margin over Symbol Recognition, but with a rather definite weakness in the Visual-Motor Association ability. As a group, it would appear that the experiences which emphasized the ability to relate meaningful visual symbols were not only somewhat diminished in school (as compared with Groups A and B for example) but their initial ability level was somewhat below par.

The gain made by Group D over the 9 month evaluation period was about 21 months for the Picture Vocabulary subtest. At the final evaluation point, this group exceeded norm standards for an "average" youngster by about 12 months. In the area of Visual Decoding, this same group gained 19 months, but remained about 7 months below the norms for children of similar chronological age.

As indicated previously, the relative area of deficit was in the Visual-Motor Association area. Gains of 9 months over the 9 month evaluation period was shown by this group. However, at the final evaluation point, Group D remained 12 months below the expectations of an "average" child of similar age.

Finally, the Group E profile shows, with one exception, a fairly well developed and even profile for the comprehension category. The exception being the dramatically high score on the Auditory Decoding subtest. This strongly suggests that this small town-rural non-disadvantaged group has acquired a fairly high ability to comprehend the spoken word and the associated skills. The visual skills are relatively lower (by 2 standard score values) than the other areas, but certainly cannot be strictly interpreted as a weakness. The Visual Decoding final score was at the "average" level; the Visual-Motor Association skill was about 6 months below the norms for a group of similar age.

The third set of graphs, Figure 23, present the results of instruments in the Quantitative Skills category. In this category, the following instruments and techniques were used: (1) Size and Number - Pictorial Test of Intelligence; and (2) Elkind Measurement of Quantitative Comparisons (a) Conservation of Number, (b) Conservation of Volume (2 parts), and (c) Conservation of Shape. The Size and Number subtest, which is a fairly global measure of quantitative skills and knowledges, spreads out the five groups in the same general order as the previously discussed tests with all means except these of Group C and D being over two standard errors apart.

The Group A profile is moderately even except for the Size and Number performance. This group gained 9 months over the 9 month evaluation period in the concepts tapped by this test. At the final evaluation point, Group A remained 15 months below the level expected for a group of this age according to the test norms. Of the three Elkind measures, a slight advantage is given to the conservation of volume. Note that this group surpasses all groups on this task by a small margin. Slight decrements are noted in the conservation of quantity and volume (shape), but for the most part, Group A "held their own" on the conservation tasks.

Group B shows a significant deficit on the Size and Number skills. A gain of only 5 months occurred over the 6 month period while the other groups were gaining 9-15 months. At the final evaluation point, Group B falls 19 months below the test norms. One can only hypothesize that a minimal amount of experience in the conceptual skills related to size and numbers was given to this group in the kindergarten year.

For 2 of the 3 conservation tasks, Group B assumes the middle position. A most slight advantage is seen for the conservation of volume tasks over the conservation of quantity task.

The profile for Group C is the most stable of all the groups. No real weaknesses are identified by the standard score method of presentation. Slight advantages are shown on the Size and Number and Conservation of Volume (shape) tasks. Although this group made a gain of 12 months over the 6 month evaluation period, they remained 6 months below the level expected for a group of this age. So, even though the profile shows no problem in Size and Number concepts in relation to the rest of the groups, a mild decrement remains for the group at the end of the kindergarten year.

The profile for Group D shows a definite strength in the Size and Number concepts and a discernible decrement in the conservation tasks (lowest of the 5 groups in 2 of 3 tasks). A gain of 12 months was shown on the Size and Number subtest over the 9 month evaluation period. However, at the final evaluation point, Group D scored about 6 months below the norms for children of similar chronological age.

This is a good example of where the profile identifies an area as a strength on the profile and yet is below the expected level when compared to the appropriate age norms of a standardized test (a similar occurrence is reported for Group C in the previous paragraph).

Lastly, Group E shows a definite strength, relatively, on the Size and Number concepts. This group gained 15 months over the 6 month evaluation period and scored 3 months above test norms at the final evaluation period. A second area of strength appears to be the conservation of quantity. There is a definite departure in an upward direction of Group E's performance from the other four groups. The two conservation of volume tasks are only slightly less developed than the other skills in this category.

Figure 24 presents the standardized mean scores for the General Ability Category. The General Ability Category included the following instruments: (1) General Information, (2) Pintner-Cunningham, (3) Caldwell, and (4) Wechsler PreSchool and Primary Scale of Intelligence - WPPSI (verbal and performance).

The profile similarities of Groups C, D and E and Groups A and B are striking. On these global measures of general ability, with one exception (Pinter-Cunningham), Groups C, D and E consistently exceed Groups A and B performance by three or more standard score units. Thus is repeated a familiar pattern shown on most of the measures in this study. Group E consistently scores about two standard score units above Groups C and D on all measures. The differences between the latter groups are non-interpretable, practically speaking. Group A exceed the performance of Group B by two standard score units on two of the five comparisons (General Information and Caldwell). The remaining three comparisons between these two groups are non-interpretable.

Figure 25 presents the standardized mean scores for two fo the affective measures. This Affective Trait cateogry included: (1) Illinois Index of Self-Derrogation, and (2) Delay of Gratification.

Again, the familiar pattern of Groups C, D and E and Groups A and B forming separate clusters is found on the Illinois Index of Self-Derrogation. Groups A and B are more prone to make negative statements about themselves than the other three groups. The implication may be that the "self-concept" of these two groups is less adequately developed than the other groups. An alternate hypothesis is that Groups A and B are far more candid than the other groups about admitting negative statements about themselves. No differences exist between the Head Start and non-Head Start urban-disadvantaged groups on the self-derrogation measure.

The Delay of Gratification meausre is difficult to interpret. The commonality that exists between Groups A, B, E and D is not clear, but they clearly choose the alternative that implied less "delay" than Group C. Previous research in this area would suggest that Groups C and E would demonstrate greater delay of gratification with Groups A, B and D being more prone to select the immediate gratification alternative. The data does not offer any firm interpretations as to the reasons for the lack of concurrence with previous work.

Summary

This study developed from the following rationale. Existing descriptions of disadvantaged children and contrasts with their more advantaged peers tend to be limited to general or global type measures which indicate quantitative deficits in general cognitive functioning but fail to point out the nature of these differences in sufficient detail to guide teachers in building curricular experiences to remediate them. Further, this detailed information is needed for a single group of children so that inter-test or inter-ability can be made. Since adequate norms do not exist for many of the measures contemplated, several control or comparison groups are needed to adequately interpret the results for the groups of primary interest - disadvantaged children attending Head Start.

While the study began with a set on the part of the investigators that a breakdown of cognitive and related affective traits into sufficiently detailed subparts would result in finding some traits in which non-disadvantaged youngsters would do relatively well and others on which disadvantaged children would do relatively well, in general, the study did not support this. The disadvantaged groups of children, particularly those in the urban setting, performed at lower levels in all areas of cognitive functioning that were measured. There were intra-group differences indicating that the seriousness of the deficits or disadvantages under which these children approach formal school instruction varies with the trait being studied. These differences have been discussed earlier in the chapter. If the primary purpose of Head Start and similar programs is to prepare children to cope more successfully with the existing expectations of our educational institutions, it becomes evident that a general enrichment program or philosophy is not likely to be sufficient. Areas of relatively more serious deficit must be identified and adult-determined, structured learning tasks, sequenced to build or improve skills and abilities in these deficit areas must be used to supplement the more general, child centered enrichment aspects of the programs.

Effects attributable to Head Start are difficult to identify. As has been mentioned, the attempt to use an urban, disadvantaged, non-Head Start control group provides inconclusive results because of the apparent finding that disadvantaged children who did not voluntarily attend Head Start are, in general, lower performing, less able and therefore more disadvantaged than the children who did attend Head Start. Looking at the curves over time and noting the similarity in shape and amount of curvature leads to the generalization that summer Head Start merely moves the starting of school ahead two to three months. The types of changes which might be attributed to the Head Start programs seem to be those which occur during the first weeks of any formal school program. The question of whether gains attributable to these extra weeks of schooling continue or whether they are lost over time cannot be answered with any certainity. On certain measures initial differences seemed to dissipate over the kindergarten year, on others the differences were accentuated by the added age and school experience.

In a more specific vein, the following conclusions and suggestions are offered.

The relatively high dependency of urban disadvantaged children on siblings (Family Relations test results) suggests additional study of utilization of older children, not necessarily siblings, for authority roles typically filled by adults.

The improvement of self-concept (Illinois Index of Self-Derrogation results) among the small town-rural disadvantaged group during the kindergarten year compared to the little change in the relatively poor self-concept of the urban disadvantaged groups could be attributed to the integration of disadvantaged with non-disadvantaged which occurred with Groups D and E. While the study was not structured to test this hypothesis, this certainly appears to be a reasonable explanation.

The performances of the urban non-disadvantaged control group, Group C, emphasizes a problem which must be faced in the allotment of monies for compensatory educational experiences for disadvantaged children. While on several measures the group's performance is comparable to that of Groups D and E, on others it is clearly of a lower level indicating substantial areas of deficiency. Present formulas for determining eligibility for Head Start and other compensatory programs on parents' income or other economic measures effectively exclude this type of child from programs he or she badly needs.

Related to this point is a generalization about Group D, the small town-rural Head Start group. When one notes their relatively strong performance across the entire battery, particularly relative to that of Group C, it becomes more apparent that economic criteria provide an inadequate basis for allocation of funds designed to equalize the educational opportunities of children. The educational disadvantages resulting from poverty vary with the geographical and cultural setting and dictate that programs and priorities for funding of programs focus more on the distinct needs of diverse groups.

REFERENCES

Anastasi, Anne. Culture - Fair Testing. *Education Digest*, 30: 9-11, April 1965.

Anastasi, Anne. Psychological Tests; Uses and Abuses. *Teachers College Record Book*, 389-93, February 1961.

Ausubel, D. P. and Ausubel, P. Ego Development Among Segregated Negro Children. In Passow, A. H. (Ed.) *Education In Depressed Areas*, New York: Teachers College, Columbia University, 1963.

Bandura, A. and Walters, R. *Social Learning and Personality Development*, New York: Holt, Rinehart and Winston, 1963.

Bean, Kenneth L. Negro Responses to Verbal and Non-Verbal Test Materials. *Journal of Psychology*, 13:343-53, April 1942.

Bene, Eva and Anthony James. *The Family Relations Test*, London: National Foundation for Educational Research in England and Wales, 1957.

Boger, T. H. An Experimental Study of the Effects of Perceptual Training on Group IQ Scores in Elementary Pupils in Rural Ungraded Schools. *Journal of Educational Research 46*, 45-53, 1952.

Brazziel, W. F. and Terrill, M. An Experiment in the Development of Readiness in a Culturally Disadvantaged Group of First-Grade Children. *Journal of Negro Education*, 31, 4-7, 1962.

Brofenbrenner, U. The Changing American Child - A Speculative Analysis. *Merrill - Palmer Quarterly 7*, (2), 73-84, 1961.

Burgemeister, B., Blum, L., and Lorge, I. *Columbia Mental Maturity Scale Revised Edition*, New York: Harcourt, Brace and World, 1959.

Buros, Oscar. *Mental Measurements Yearbook (6th)*, Highland Park, N. J.: The Gryphon Press, 1965.

Casler, L. Maternal Deprivation: A Critical Review of the Literature. *Society for Research in Child Development Monograph*, 26, No. 2, 1961.

Clarke, A. D. B., and Clarke, A. M. How Constant is the IQ? *Lancet* 265:877-80, October 24, 1953.

Cloward, R. A. and Jones, J. A. Social Class: Educational Attitudes and Participation. In Passow (Ed.) *Education in Depressed Areas*, New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Covington, M. V. Some Effects of Stimulus Familiarization on Discrimination. Unpublished Doctoral Dissertation, University of California, reported in Bloom, Davis, Hess. *Compensatory Education for Cultural Deprivation*, New York: Holt, Rinehart and Winston, 1965.

Davidson, Kenneth S. and Others. A Preliminary Study of Negro and White Differences on Form I of the Wechsler Bellevue Scale. *Journal of Consulting Psychology*, 14:489-92, October, 1950.

Davis, Allison and EElls, Kenneth. *Manual for the Davis-Eells Test of General Intelligence or Problem Solving Ability*, Yonkers-on-the Hudson, New York: World Book Company, 1953.

Deutsch, C. P. Auditory Discrimination and Learning: Social Factors. *Merrill - Palmer Quarterly* 10, 277-96, July, 1964.

Deutsch, Martin. The Role of Social Class in Language Development and Cognition. *American Journal of Orthopsychiatry* 35, 78-88, January, 1965.

Deutsch, M. and Brown, B. Social Influence in Negro-White Intelligence Differences. *Journal of Social Issues*, 20, 24-35, April, 1964.

Deutsch, Martin and Others. Guidelines for Testing Minority Group Children. *Journal of Social Issues*, 20:129-45, April, 1964.

Deutsch, M. Early Social Environment: Its Influence on Social Adaption. In Schreiber (Ed.) *The School Dropout*, Washington, D. C.: The National Education Association, 1964.

Douglas, J. W. B. *The Home and The School*, London: McGibbon and Kee, 1964.

Dreger, R. M. and Miller, K. Comparative Psychological Studies of Negroes and Whites in the United States. *Psychological Bulletin*, 57:361-402, 1960.

Edwards, T. B. and Webster, S. W. Correlates and Effects of Ethnic Group Identification. *Research Relating to Children*, Children's Bureau, Bulletin No. 17, Washington, D. C.: Government Printing Office, 1963.

3

Elkind, D., The Development of Quantitative Thinking: A Systematic Duplication of Piaget's Studies, *Journal of Genetic Psychology*, 1961, 98:37-46.

French, J. L. *Pictorial Test of Intelligence*, Boston: Houghton-Mifflin, 1964.

Frostig, M., et.al. MARIANNE FROSTIG Developmental Test of Visual Perception, Third Edition, Palo Alto, California Consulting Psychologist Press, Inc., 1964.

Goff, R. M. Some Educational Implications of the Influence of Rejection on Aspiration Level of Minority Group Children. Journal of Exceptional Education, 23:179-183, 1954.

Gordon, E. W. A Question of Culture. American Child, 45:11-14, 1963.

Gordon, E. W. Characteristics of Socially Disadvantaged Children. Review of Educational Research, 35:377-388, 1965.

Higgins, Conwell and Sivers, Cathryne H. A Comparison of Stanford-Binet and Colored Raven Progressive Matrices IQ's for Children With Low Socio-economic Status. Journal of Consulting Psychology, 22:465-68, December, 1958.

Horrocks, John E. Assessment of Behavior, Columbus: Chas. E. Merrill Books, Inc., 1964.

Hunt, Joseph McV. Intelligence and Experience, New York: Ronald Press Co., 1961.

Irwin, O. C. Infant Speech: The Effect of Family Occupational Status and of Age on Use of Speech Types. Journal of Speech and Hearing Disorders, 13:224-26, September, 1948.

John, V. P. The Intellectual Development of Slum Children: Some Preliminary Findings. American Journal of Orthopsychiatry, 33:813-822, 1963.

Kagen, J., et.al. Information Processing in the Child: Significance of Analytical and Reflective Attitudes. Psychological Monograph (1964), 78, Whole No. 578.

Kagen, J., Moss, H. A., and Sigel, I. E. Psychological Significance of Styles of Conceptualization. In J. E. Wright and J. Kagen (Eds.) Basic Cognitive Process in Children, Monograph, Society for Research in Child Development, 28, No. 2, 1963.

Karp, Joan and Sigel, Irving. Psychoeducational Appraisal of Disadvantaged Children. Review of Educational Research, 35, 401-402, 1965.

Kidd, Aline H. The Culture - Fair Aspects of Cattell's Test of g: Culture-Free. Journal of Genetic Psychology, 101:343-62, December, 1962.

Klineberg, Otto. Negro-White Differences in Intelligence Test Performance: A New Look at an Old Problem. American Psychologist, 18:198-203, April, 1963.

Kohn, M. L. Social Class and the Exercise of Parental Authority. American Sociological Review, 24:352-66, 1959.

Larson, Richard and Olson, James L. A Method of Identifying Culturally Deprived Kindergarten Children. Exceptional Children, 30:30-34, November, 1963.

Lehmann, Irvin J. Rural-Urban Differences in Intelligence. Journal of Educational Research, 53:62-68, October, 1959.

LeShan, L. L. Time Orientation and Social Class. Journal of Abnormal and Social Psychology, 47:589-592, 1952.

Lesser, Fifer and Clark. Mental Abilities of Children in Different Social and Cultural Groups, U. S. Department of Health, Education and Welfare, Office of Education, Cooperative Research Project No. 1935. Washington, D. C.: Government Printing Office, 1964.

MacArthur, R. S. and Elley, W. B. The Relocation of Socio-Economic Bias in Intelligence Testing, British Journal of Educational Psychology, 33:107-19, June, 1963.

McCarthy, J. J. and Kirk, S. A. Illinois Test of Psycholinguistic Abilities, Experimental Edition, Urbana, Illinois: University of Illinois Press, 1963.

Meyerwitz, J. H. Self-Derrogations in Young Retardates and Special Class Placement. Child Development, 1962, 33:443-451.

Mischel, W. Father-Absence and Delay of Gratification: Cross-Cultural Comparison. Journal of Abnormal and Social Psychology, 1961, 63:116-124.

Neugarten, B. L. Social Class and Friendships Among School Children. American Journal of Sociology, 51:305-313, 1946.

Newton, E. S. Planning for the Language Development of Disadvantaged Children and Youth. Journal of Negro Education, 33:264-74, Summer, 1964.

Noll, Victor H. Relation of Scores on Davis-Eells Games to Socio-Economic Status, Intelligence Test Results, and School Achievement. Educational and Psychological Measurement, 20:119-29, Spring, 1960.

Office of Economic Opportunity, *Project Head Start, Daily Program for a Child Development Center*, Washington, D. C., Office of Economic Opportunity, 1965.

Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. *The Measurement of Meaning*, Urbana, Illinois: University of Illinois Press, 1957.

Pasamanich, Benjamin and Knobloch, Hilda. Early Language Behavior in Negro Children and the Testing of Intelligence. *Journal of Abnormal and Social Psychology*, 50:401-402, May, 1955.

Piaget, J. and Inhelder, Barbel. Diagnosis of Mental Operations and Theory of Intelligence. *American Journal of Mental Deficiencies*, 51:401-406, 1947.

Porteus, S. D. *The Porteus Maze Test and Intelligence*, Palo Alto: Pacific Books, 1950.

Riessman, Frank. *The Culturally Deprived Child*, New York: Harper and Row, 1962.

Riessman, F. The Culturally Deprived Child: A New View. *Programs for the Educationally Disadvantaged*, U. S. Department of Health, Education and Welfare, Office of Education, Bulletin, 1963, No. 17. Washington, D. C.: Government Printing Office, 1963.

Sievers, D. J. *Development and Standardization of a Test of Psycholinguistic Growth in Preschool Children*, Doctor's Thesis, Chicago: University of Illinois, 1955. (Abstract: Dissertation Abstracts 16:286-287, No. 1, 1956.)

Sigel, Irving E. The Attainment of Concepts. In Hoffman and Hoffman (Eds.) *Review of Child Development Research*, Volume 1, New York: Russell Sage Foundation, 1964.

Siller, Jerome. Socioeconomic Status and Conceptual Thinking. *Journal of Abnormal and Social Psychology*, 55:365-71, November, 1951.

Stendler, C. B. Social Class Differences in Parental Attitudes Toward School at Grade I Level. *Child Development*, 22:37-47, 1951.

Wechsler, David. *Wechsler Preschool and Primary Scale of Intelligence*, New York: The Psychological Corporation, 1967.

Wentland, T. J. and Geeding, W. C. A Descriptive Study of the WISC Pattern of Five-Hundred Disabled Readers. Paper presented to Wisconsin Educational Research Association, 1967.

Whiteman, Martin. Intelligence and Learning. *Merrill - Palmer Quarterly* 10:297-301, July, 1964.

APPENDICES

APPENDIX A

This appendix contains descriptions of the instruments and techniques used in this study. They are presented within the classification framework and in the order in which they are discussed in Chapter IV. This classification and order is:

A. Ability to make discriminations.
  1. Columbia Mental Maturity Scale
     Frostig Developmental Test of Visual Perception
  2. Figure-Ground subtest
  3. Spatial Relations subtest
  4. Haptic Visual Matching Test

B. Basic comprehension skills.
     French Pictorial Test of Intelligence
  5. Picture Vocabulary subtest
     Illinois Test of Psycholinguistic Abilities
  6. Visual Decoding subtest
  7. Visual-Motor Association subtest
  8. Auditory Decoding subtest
  9. Symbol Recognition Test, Parts I and II

C. Basic quantitative skills.
     French Pictorial Test of Intelligence
  10. Size and Number subtest
  11. Elkind Measurement of Quantitative Comparisons

D. General ability and knowledge.

    12. General Information Test

    13. Pintner-Cunningham Group Intelligence Test

    14. Caldwell Preschool Inventory

    15. Wechsler Pre-School and Primary Scale of Intelligence

E. Cognitive style.

    16. Block Sorting Task

    17. Matching Familiar Figures

F. Affective traits.

    18. Illinois Index of Self-Derrogation

    19. Delay of Gratification

G. Family relationships

    20. Family Relations Test

1. Columbia Mental Maturity Scale

The Columbia Mental Maturity Scale, Revised Edition authored by Bessie B. Burgemeister, Lucille Hollander Blum, and Irving Lorge, was published in 1959 by Harcourt, Brace and World, Inc. It can best be described as a pictorial type classification test. It consists of 100 items presented on heavy cards which measures 6 x 19 inches. Each card has three to five drawings.

Although originally designed as a measure of intelligence for handicapped children, the revised edition is viewed as more appropriage as a measure of learning aptitude of nondisabled, rather than disabled, children. (Buros, 1965)

The Columbia was selected primarily to describe the discrimination abilities of children rather than to provide an estimate of I.Q. The major advantage of the instrument is that is requires no verbal response and a minimum of motor responses which seemed desirable given the age and socio-cultural background of the groups being studied. The test is untimed, individually administered and requires perceptual discriminations involving color, shape, size, use, number, kind, missing parts, and symbolic material. The subject selects the picture in each series which is different from, or unrelated to, the others. The discrimination required is that of recognizing the picture which does not belong with the others, i.e., the education of a principle for organizing the pictures so as to exclude just one. The items are arranged in order of difficulty.

The revision of the test from its original form reduced the sampling of acculturation factors (less reading, less demand on geographic and historical information).

Studies which have reported correlations between the Columbia and other intelligence measures yield, in summary, the following:

Correlations with	Age	N	r
Stanford Binet (1937)	3 to 12	957	.78
Stanford Binet (1960)	4 and 5	not reported	.39
WISC Verbal	2nd grade	not reported	.45
WISC Performance	2nd grade	not reported	.52

Reliabilities reported in the manual are obtained from split-half, Spearman-Brown estimates. Based on samples of 79 to 139 children the following are reported: age 4, r = .89; age 5, r = .91; age 9, r = .91; age 10, r = .92.

2. & 3. Frostig Developmental Test of Visual Perception

Two subtests of the Frostig Developmental Test of Visual Perception were selected as the primary means of identifying the degree of differential perceptual development between groups. The Frostig is composed of five subtests selected on the basis of their apparent relevance to performance in nursery school, kindergarten, and the primary years, particularly those behaviors related to reading readiness. The subtests selected for use and described below were Figure-Ground and Spatial Relations.

Limited technical data is available on the Frostig. The manual provides a table of inter-subtest correlations based on 385 kindergarten age children. Spatial Relations and Figure-Ground correlated .37, while Spatial Relations correlated .71 with Total Score compared to a correlation of .73 between Figure-Ground and Total Score.

Two reliability studies are reported with children of the age range included in this project. Test-retest reliability with a two week interval was obtained on 55 kindergarten children. Figure-Ground had a reliability coefficient of .46 and Spatial Relations a coefficient of .66. Using a procedure based upon matching subtest items on the basis of difficulty, a split-half reliability study was conducted at various age levels. In a 5-6 year old group of 364 children, split-half reliability coefficients of .93 for Figure-Ground and .85 for Spatial Relations were obtained.

The Figure-Ground subtest involves shifts in perception of figure against increasingly complex grounds. Intersecting and hidden geometric forms are used as stimuli and the subject is asked to discriminate these figures from the ground in which they are embedded. This subtest requires recognition only. The author maintains that sufficient ability to shift in figure-ground perception is necessary for the analysis and synthesis of words, phrases, sentences, and paragraphs involved in reading. The rationale for the subtest is based upon the author's conclusion that children who could not recognize words often seemed to have disturbances in figure-ground perception.

The Spatial Relation subtest involves the analysis of simple forms and patterns. These consist of lines of various lengths and angles which the child is required to copy using dots as guide points. Accurate perception of spatial relations is seen by the author as the basis of the ability to differentiate similar letters like "b" and "d", and to recognize the sequence of letters in a word or of words in a sentence. Children with poor perception of spatial relations tend to interchange the order of letters in words and can neither read nor spell longer words.

The Frostig scoring procedures for these subtests are basically an all or nothing system. The child very closely approximating the complete task receives one point, if not, zero. Since the children in this study were somewhat younger than the group on which the Frostig was normed and since there was reason to believe the children being studied would, over all, have less than normal percpeutal development, the investigators decided to revise the scoring procedures allowing credit for partial scores.

The procedures followed was essentially one of subdividing each task into equally weighted subtasks. Weights for subtests were such that the total points for all single items equalled ten. The Frostig manual suggests that points be subtracted for errors. In the Figure-Ground subtest three types of error designated as deviancy, extraneous responses and closure are identified. For the Spatial Relations subtest deviancy and placement errors require attention. Deviancy errors result from the child drawing response lines which are in the appropriate area but deviate from those suggested by the stimulus materials. Extraneous errors are the result of drawing extra lines, while closure errors are failures to properly bring intersecting lines together. Drawing the response in the wrong area of the matrix is termed a placement error.

These types of errors were identified and then categorized as major or minor errors. For major errors of any one type a subject's total score for an item was reduced 20% and for a minor error it was reduced to 10%. Thus, on the Figure-Ground subtest each item was scored by first counting the number of subparts completed in the task and then this number was reduced by from 0 to 60% depending upon the presence and relative severity of errors of the three types. A similar procedure was used with the Spatial Relations test except that errors could only reduce a score a maximum of 40%.

While there are some problems with using this cumbersome a scoring scheme in a clinical situation with an individual subject, this was a somewhat different situation. The purpose was to describe and differentiate among groups using group means. For this purpose we did obtain stable and apparently meaningful results. A pilot study of the scoring procedure produced inter-score correlations greater than .90 on a sample of 30 protocals.

4. Haptic Visual Matching Test

This technique and one described in the Cognitive Style category, Matching Familiar Figures, represent an attempt to measure the impulsivity-reflectivity dimension of children's cognitive styles as well as being a discrimination task. Both procedures are derived from the work of Kogan, Moss and Sigel (1963). The rationale for the category of cognitive style is more completely developed in Section 17, Matching Familar Figures, of this Appendix.

The Haptic task allows the child to first explore with his fingers a wooden form to which he had no visual access. Subjects were allowed unlimited time to explore the form and then were presented with a visual array of stimuli, one of which replicated the form that had been explored haptically.

As the children in this study were somewhat younger than those studied by Kogan, Moss and Sigel, the procedure required modification. The device developed for conducting the test was a large wooden box, approximately two feet square. That side of the box placed facing the child was open, but covered by a heavy cordoroy cloth. In the cover were two slits encircled by elastic. The subject would place his hands into the slits and handle the stimulus object placed in his hand by the examiner. The back of the box was hinged to allow the examiner to hand stimuli to the child. When the subject finished manipulating the stimulus, he dropped the object, withdrew one or both hands and made a choice by pointing from an array of four response choices presented to him. This array was on a single sheet of 8 x 11 paper in a looseleaf holder.

While the original instrument had 20 items, the modified procedure used only 16. In choosing items to retain and in building new items complexity was reduced and physical features were accentuated.

Two scores were obtained. The first was the number of correct responses and the second of the time, in seconds, from the first touching of the stimulus object to completion of the response choice.

5. Pictorial Test of Intelligence; Picture Vocabulary

This test was developed by Joseph L. French to provide an easily administered, objectively scored, individual testing instrument to be used in assessing the general intellectual level of both normal and handicapped children between the ages of three and eight. To respond to the items of this test, children who understand English need only be capable of hearing simple verbal instructions and responding to visual stimulation. Subjects indicate their responses to questions by pointing to pictorial symbols of their choice on large response cards.

According to the author the test is, in some respects, a further development of Binet type scales, but employs an objective, multiple-choice technique.

For the present study two of the six subtests were utilized, Picture Vocabulary and Size and Number (described below in Section 10). The purpose of the Picture Vocabulary subtest is to measure verbal comprehension. The items require recall of previously acquired verbal meanings. Children respond to a word spoken by the examiner and then select, from the response card, that one of four drawings which best represents a meaning of the stimulus word. The subtest, containing 32 items, is terminated after six consecutive errors.

The author states that choice of item words is based upon frequency of appearance in texts and supplementary books used in lower grades and upon observation of children's word usage in a wide variety of situations.

Reliability data for the subtests are not provided in the manual. Reported total test reliabilities for ages 4, 5 and 6 are .90, .92 and .93 respectively. The corresponding standard errors of measurement are 3.49, 3.29 and 2.56. Test-retest correlations over periods of two to six weeks are reported in the range from .90 to .96 but these are based on samples of, at most, 31 subjects.

Correlations with the WISC are reported (sample size is only 32);

.38 with Verbal, .33 with Performance and .38 with the Total. Means and standard deviations on the Picture Vocabulary subtest for ages 4, 5 and 6 are:

Age	Mean	Standard Deviation
4	12.53	4.67
5	15.53	4.84
6	19.05	4.01

6., 7. & 8. The Illinois Test of Psycholinguistic Abilities; Visual Decoding, Visual-Motor Association and Auditory Decoding.

The version of this test used in this study was the experimental edition. The battery was designed by its authors, J.J. McCarthy and S.A. Kirk, to meet the need for a comprehensive instrument for the assessment of language development in exceptional children, particularly those of preschool age.

The theoretical model which gives rise to the rationale of the subtests comprising the Illinois Test of Psycholinguistic Abilities is Osgood's psycholinguistic model (1957a, 1957b). Although the ITPA contains nine subtests, of which only three were utilized in this study, it is advisable to briefly describe the basic theory and total battery in its entirety.

The theory postulates three major dimensions of psycholinguistic abilities; level of organization, psycholinguistic processes, and channels of communication. Each of these major dimensions are further subdivided.

The level of organization refers to the functional complexity of the organism. Osgood has identified two levels of organization as relevant to the acquisition of language. The *representational* level is reached when the organism is sufficiently organized to mediate activities requiring the meaning or significance of linguistic symbols. The *automatic-sequential* level refers to the organism being able to mediate activities which require the retention of linguistic symbol sequences and the execution of automatic habit-chains.

The second major dimension, psycholinguistic processes, refers to the habits required for normal language usage and are categorized into three main subsets. *Decoding* refers to the total of those habits required to ultimately obtain meaning from either visual or auditory linguistic stimuli. *Encoding* refers to the total of those habits required to ultimately express oneself in words or gestures. *Association* refers to the total of those habits required to manipulate linguistic symbols internally.

The final major dimension, channels of communication, describes the sensory-motor path over which linguistic symbols are received and responded to. It is divided into mode of reception, auditory or visual, and mode of response, vocal or motor. Thus from an assessment frame of reference, to test pure decoding ability only the mode of reception need be specified, to test encoding only the mode of response, while to test association ability, the entire channel must be specified, e.g., Auditory-Vocal Association, Visual-Motor Sequencing, etc.

The three subtests administered as part of this study, Visual Decoding, Visual-Motor Association, and Auditory Decoding, are all at the representational level as opposed to the automatic-sequencing level.

Internal consistency reliability estimates on children in the 4.5 to 6.5 year age range are reported as being between .69 and .84 for the Visual Decoding subtest, between .70 and .83 for Visual-Motor Association and between .84 and .89 for Auditory Decoding. While intercorrelations are reported among the three subtests, none are large enough to be significantly different than zero.

The Visual Decoding subtest measures the ability of the subject to comprehend pictures and written words. It is assessed by a picture identification technique in which the subject selects from among a set of pictures the one which is most nearly identical, on a meaningful basis, to a previously exposed stimulus picture. The test, consisting of twenty-four items of increasing difficulty, is terminated when the subject fails six of eight consecutive items.

The Visual-Motor Association subtest measures the ability of the subject to relate meaningful visual symbols. The emphasis is more upon a functional or thematic relationship between stimulus and response, in contrast to the categorical relationship assessed in the Visual Decoding subtest.

The subtest begins with the use of miniature objects of life-size small objects. The subject is asked to make an association between one of four objects and the other three. The subject responds by pointing to the correct answer. The first four such items use concrete objects as stimuli. The balance of the items contain a stimulus picture and a comparison picture with four choices. There are twenty-eight items and the test is terminated when three of four consecutive items are missed.

The Auditory Decoding subtest assesses the ability of the child to comprehend the spoken word. This ability is measured by a controlled vocabulary test in which the subject is asked to answer yes or no by voice to a series of graded questions. Examples of the item questions are, Do you eat? Do airplanes fly? Do cars cry?, etc. There are thirty-six items arranged in increasing difficulty as to the vocabulary levels of the words used. The test is terminated when four of eight consecutive items are failed.

9. Symbol Recognition Test

The Symbol Recognition Test developed by Richard Larson and James Olson, is an experimental procedure designed to measure the ability of children to detect subtle differences among symbols of familiar objects which is seen as a basic prerequisite for the development of conceptualizing skills and the use of verbal symbols in language and reading development. The test is administered in two parts, each part consisting of thirty-one items. Each item contains three figures. The first part contains stylistic symbolizations of objects. The subject is asked to identify a specific object from the three symbols, e.g., turtle, ship, tree, church, etc. The second part of the test contains an exact duplication of the same three objects on all items, but each is pictured in a realistic or complete way. In other words there is an actual sketch of a turtle, a tree, etc. The three figures comprising an item in Part I are such that there is one basic symbol which is common to all three figures, but each contains some symbolic characteristic that is the clue to its identity.

Part II is intended to facilitate the interpretation of the results of Part I by differentiating those errors which are a function of inability to recognize the symbolic representations of objects from which are a function of unfamiliarity with the objects.

10. Pictorial Test of Intelligence; Size and Number

The purpose of the Size and Number subtest of the French PTI is to sample the child's ability in the areas of perception and recognition of size, number symbol recognition and comprehension, ability to count, and ability to solve arithmetical problems. In a number of the items the subject must employ reasoning with numerical concepts.

Problems of basic facts (addition, subtraction, multiplication, division) and fractions are included, but recognition of number symbols is not necessary for their solution. Because the items are arranged in approximate order of difficulty and six consecutive failures terminates testing, few of the arithmetic problem type of items were actually given.

As indicated above in Section 5, the manual reports no reliability data by subtest. Correlations between this subtest and the Verbal, Performance and Total scales of the WISC are .53, .42 and .50. For children in the age range studied here means and standard deviations are reported.

Age	Mean	Standard Deviation
4	6.73	3.87
5	10.59	2.79
6	16.21	4.97

11. Elkind Measurement of Quantitative Comparisons

This technique attempts to assess the developmental level of the subjects' quantitative thinking based upon the theoretical rationale of Piaget. The specific procedures used are modifications of Elkind's replications of some of Piaget's studies. (Elkind, 1961) Elkind's work was an attempt to apply statistical methods and systematic design to the clinical methods preferred and used by Piaget.

> "Piaget assumes that the child's quantitative thinking develops and that the child's success in comparing quantity earmarks this development. By studying the child's responses to quantity comparisons of increasing difficulty Piaget sought to trace the developmental changes in the form and content of the child's quantitative thought.
>
> In his work Piaget distinguishes three types of perceived quantity by which things can be compared without actual measurement. The simplest type of perceived quantity he speaks of as gross. Gross quantities are single perceived relations between objects (longer than, larger than) which are not coordinated with each other. A more complex type of quantity Piaget calls intensive. Intensive quantities are perceived quantity relations taken two by two (longer and wider, taller and thicker). The most complex type of perceived quantity Piaget calls extensive. Extensive quantities are unit relations between objects (x is half of y, x is twice y, etc.). Extensive quantities are logical constructs which must be attained by abstraction or reasoning but Piaget assumes that once attained, the subject perceives them directly as perceptually given properties of the object." (Elkind, 1961)

The basic rationale of the tasks employed is the comparison of two amounts of material several times. The two amounts are successively arranged so that sometimes a comparison of gross, sometimes a comparison of intensive, and sometimes a comparison of extensive quantity is the minimum requirement for a successful result. Three types of materials used for making the comparisons are liquids, sticks, and beads. The task is administered in two sessions, one to three days apart. In the first session the sticks, identified as candies, are used to test for conservation of number and the orange colored liquid, identified as Kool Aid, to test for conservation of volume. In the second setting the sticks, identified as trains, and the wooden beads are used.

In each of the four sequences, the subject is required to make six responses, either verbal or motor. Summing over both sessions results in a total score between 0 and 12 for both the quantity and volume conservation tasks. One item in each quantity sequence and two items in each volume sequence required extensive comparisons. Scores on these three types of extensive comparison items (either 0, 1 or 2) are also reported.

12. General Information Test

The General Information Test is a short test of the child's general knowledge of his environment developed by Richard Larson and James Olson. It consists of twenty short answer questions that deal with the child's general knowledge about his immediate environment and the types of knowledge that children acquire through exposure to the general type of pre-school literature.

Examples of questions from the instrument are:

a. What city do you live in?

b. What street do you live on?

c. What month is this?

d. What is the coldest time of the year?

e. Where do bears live?

Although some of the answers fall into the right or wrong category, the investigators modified the scoring procedure to allow for scoring a response as 0, 1, or 2. No credit was given for a completely inaccurate or wrong response, or for no response at all. One point was given for those responses that indicated some knowledge relevant to the item content but not a precise, accurate, completely satisfactory response. A full response received two points.

13. Pintner-Cunningham Primary Test of General Abilities (Revised)

This test was used because it is administered to the Milwaukee kindergarten students in their second semester of kindergarten as a part of the city-wide testing program. It thus provided both a general I.Q. measure and a way of relating our results to the ongoing testing program of the schools. It should also be noted that this was the only procedure that was group administered.

The Pintner-Cunningham Primary Test is intended for use at kindergarten and Grade 1 levels. It is composed entirely of pictures which are marked by the subjects according to the examiner's verbal directions.

No knowledge of words or numbers is assumed. The method of indicating the answer is such that even children with poor muscular development and control are not handicapped. The arrangement of the material into a number of short subtests not only assures the cooperation of the children, but gives them a feeling of accomplishment throughout the test. Detailed directions are given for each subtest, and the time limits permit all children to do all they are able to do. Enough easy material is included so that even the slowest child is able to have some success.

The Pintner-Cunningham test is divided into seven subtests: 1) Common Observation, 2) Esthetic Differences, 3) Associated Objects, 4) Discrimination of Size, 5) Picture Parts, 6) Picture Completion, and 7) Dot Drawing. Data reported in the report are for total test raw score and I.Q. scores only.

The authors state that many other phases of intelligence might have been included, but these seven have proved to be highly discriminating at these age levels. The average three year old child finds these tests extremely difficult, while the average eight year old finds them quite easy.

The authors report a split-half reliability for total score of .84 and alternate form reliabilities comparing old and new forms ranging from .73 to .84. The standard error of measurement is 3.15.

14. Caldwell Preschool Inventory

The Preschool Inventory developed by Bettye M. Caldwell is designed to find out whether children have acquired certain skills that are ordinarily observable in children by the time they are five or six years of age. The author states that it is not a test of intelligence. The items represent a sample of some familiar types of material that are often included in a kindergarten curriculum.

This test was selected for inclusion in the study for several reasons. First, it had been used extensively in the evaluation of early Head Start programs and would therefore provide a link for comparisons between the groups of this study and other Head Start studies. Second, the Caldwell represented the only academic-oriented global measure of performance or achievement to be used. Given the conditions that the majority of research and evaluative data utilized in the public schools is typically of a global nature it seemed wise to make provision for such data in this study. While later two other global-type measures were also included, the Pintner-Cunningham and the WPPSI, the Caldwell seemed to be tapping a different type of cognitive product.

The form of the Caldwell used was the standardized edition which contains 85 items in comparison to the 148 items of the original edition. Items 1-26 have been categorized by the author as measuring Personal-Social Responsiveness. This includes four basic types of questions: (1) those dealing with self-knowledge, e.g., "What is your first name?", "How old are you?"; (2) knowledge of parts of the body, e.g., "Show me your neck", "What do we call this?" (pointing to elbow); (3) knowledge of and ability to perform specific movements, e.g., "Wiggle!", "Raise your hand!"; and (4) concepts of color, size, and spatial relations, e.g., "Put the blue car under the black box", "Put the yellow car on the little box". The second subpart, items 27-47, is labeled Associative Vocabulary. Representative samples of the items included in this section are: "Which way does an elevator go?" (child must give both verbal and motor response), "When do we eat breakfast?", "If you were sick who would you go to?", and "What does a policeman do?" Items 48-66 consitute the Concept Activation-Numerical subpart. Sample items from this section include: "How many hands do you have?",(using two groups of checkers, 2 in one and 8 in the other) "Which has more checkers in it?", (using 5 checkers in a row) "Give me the middle one". The final subpart, items 67-85, has been labeled as Concept Activation-Sensory and includes such items as: (using a page with a line, circle, square and triangle, pointing to the circle) "Make a drawing like this. Make it right here" (pointing to blank space beside circle model), (using the same stimulus sheet as above) "Which one is most like a tent?", "Which is slower, a car or a bicycle?", and (using 8 crayola crayons, pointing to red crayon) "What color is this?"

15. Wechsler Pre-School and Primary Scale of Intelligence

The Wechsler Pre-School and Primary Scale of Intelligence (WPPSI) is composed of 10 subtests, 5 verbal and 5 performance. One supplementary verbal subtest is included. The *verbal* subtests include: (1) Information, (2) Vocabulary, (3) Arithmetic, (4) Similarities, and (5) Comprehension and the supplementary subtest sentences. The *performance* subtests include: (1) Animal House, (2) Picture Completion, (3) Mazes, (4) Geometric Design, and (5) Block Design.

The WPPSI was standardized on a sample of 1200 cases. 100 boys and 100 girls in each of the six age groups, ranging by half-years from 4-6-1/2 were selected in accordance with the 1960 United States Census. In addition to the age-sex variable, geographic region, urban-rural residence, color and father's occupation approximated the percentages listed in the 1960 census. Although the test author utilized a stratified sampling technique, no specific information is presented in the manual relative to the test characteristics from the lower socio-economic classes.

Eight of the eleven WPPSI subtests are tests reincorporated from the Wechsler Intelligence Scale for Children. These eight include: (1) Information, (2) Vocabulary, (3) Arithmetic, (4) Similarities, (5) Comprehension, (6) Picture Completion, (7) Mazes, and (8) Block Design. The reader is referred to the research pertaining to the WISC for an elucidation on what factors or skills these particular subtests seem to measure. Three new tests are added to the previous eight to complete the battery. These three subtests and the skills required for performing are (1) Animal House: requires the child to associate sign with symbol. . .Wechsler states this "maybe considered as a measure of learning ability". Memory, goal awareness, attention span and ability to concentrate are all factors seemingly involved in this subtest. (2) Geometric Design: The child is required to reproduce (with pencil) selected geometric figures. These abilities depend primarily upon perceptual (visual) and visual-motor organization. (3) Sentences: The child is required to verbally repeat sentences spoken by the examiner. Probably the skills involved include auditory sequencing and memory and attention and concentration.

Technical data for the WPPSI is abundant in the manual. Split-half reliability coefficients (r) and standard errors of measurement (SE_m) for the subtests and the verbal, performance and full scale I.Q.'s for the age groups 4-6-1/2 are included. The highest and lowest percentages listed in the 1960 census. Although the test author utilized a stratified sampling technique, no specific information is presented in the manual relative to the test characteristics from the lower socio-economic classes.

Eight of the eleven WPPSI subtests are tests reincorporated from the Wechsler Intelligence Scale for Children. These eight include: (1) Information, (2) Vocabulary, (3) Arithmetic, (4) Similarities, (5) Comprehension, (6) Picture Completion, (7) Mazes, and (8) Block Design. The reader is referred to the research pertaining to the WISC for an elucidation on what factors or skills these particular subtests seem to measure. Three new tests are added to the previous eight to complete the battery. These three subtests and the skills required for performing are (1) Animal House: requires the child to associate sign with symbol. . .Wechsler states this "maybe considered as a measure of learning ability". Memory, goal awareness, attention span and ability to concentrate are all factors seemingly involved in this subtest. (2) Geometric Design: The child is required to reproduce (with pencil) selected geometric figures. These abilities depend primarily upon perceptual (visual) and visual-motor organization. (3) Sentences: The child is required to verbally repeat sentences spoken by the examiner. Probably the skills involved include auditory sequencing and memory and attention and concentration.

Technical data for the WPPSI is abundant in the manual. Split-half reliability coefficients (r) and standard errors of measurement (SE_m) for the subtests and the verbal, performance and full scale I.Q.'s for the age groups 4-6-1/2 are included. The highest and lowest reliability coefficients and the SE_m for the subtests for the age groups 5-1/2 and 6-1/2 is included below:

	Age Group 5-1/2 r	Age Group 5-1/2 SE_m		Age Group 6-1/2 r	Age Group 6-1/2 SE_m
Mazes	.91	.87	Mazes	.88	1.04
Information	.81	1.31	Information	.75	1.50
Verbal IQ	.90	4.68	Verbal IQ	.89	5.07
Performance IQ	.91	4.82	Performance IQ	.86	5.43
Full Scale IQ	.94	3.74	Full Scale IQ	.92	4.14

Across all age levels (4-6-1/2) the Mazes subtest appears the most reliable (avg. r = .87) where the split-half technique is used and Spearman-Brown correction is applied. The least reliable is the Animal House subtest (avg. r = .77).

Scoring: At each age level and for each subtest, the raw score distribution was converted to a scale with a mean of 10 and a standard deviation of 3.

Other technical data includes coefficients of correlation of WPPSI scaled scores and I.Q.'s with the I.Q.'s of three other tests. Some of this data is included below.

WPPSI	Stanford-Binet (Form L-M)	Peabody Picture Vocabulary Test	Pictorial Test of Intelligence
Verbal IQ	.76	.57	.53
Performance IQ	.56	.44	.60
Full Scale IQ	.75	.58	.64

It is most apparent that the WPPSI correlates best with the Stanford-Binet (Form L-M).

16. Block Sorting Task

This procedure was developed by Milton C. Hillery as an attempt to reduce classification, sorting or grouping behavior to its most simple form in terms of the stimuli used and consistency of attributes. Although the procedure is not identical with any other technique reported in the literature of which the authors are familiar, the principles involved have been incorporated into a large number of classification analysis procedures.

The purpose was to determine if there was a developmental hierarchy of sorting or grouping behavior based on the attributes of color, size or shape (form). Stimulus materials were presented in both 3-dimensional block form and 2-dimensional picture form.

The actual task presented subjects is that of selecting from among three alternatives that which is "most like" or "goes with" a focus stimulus. For example, subject may be presented a large red, triangular block and asked which of three alternatives - a small red circular block, a large blue square block or a small white triangular block - goes with or is like the former. The 2-dimensional materials were presented as facing pages in a loose-leaf notebook.

Sixteen items were included in each of the 3-dimensional and 2-dimensional tasks. Eight of the items were common to both parts. While second and third choices were elicited from subjects along with subjects verbalizations of reasons for the choices, only the first choice data was analyzed.

17. Matching Familiar Figures

This subtest and the Haptic Visual Matching Test described above in Section 4 are modifications of procedures developed by Kogan, Moss and Sigel (1963). Their work is an attempt to explain individual differences in intellectual behavior not adequately explained by existing theoretical and experimental work. They defined a class of variables that holds promise for explanation of this differential behavior. This set of variables is labeled as cognitive style which is defined as "stable individual preferences in mode of perceptual organization and conceptual categorization of the external environment."

Elaborating upon this Kogan (1965) has stated:

> "Three sequential operations typically occur when a person is confronted with a problem: an initial categorization of the relevant information, storage of the coded categorization, and, finally, the imposing of rules of reasoning or transformations (e.g., a complex set of verbal associations) upon the encoded data. The nature of the categorization, transformation, or elaboration mediation is governed, of course, by the nature of the problem. Students of cognitive development have generally assumed that the striking differences between the intellectual products of children of different ages or among children of the same age were attributable primarily to differences in the availability of vocabulary, possession of deductive or inductive rules, and richness of the associational network. In essence, the superior intellectual performance of older, in contrast to younger, children has been ascribed to the greater knowledge repertoire of the older children. This supposition is intuitively attractive and empirically verified. It is not surprising, therefore, that psychologists have not seriously entertained the possibility that other factors may contribute to age and individual differences in the form and quality of cognitive products. Specifically, there has been a tendency to ignore the relevance of differences in the degree of reflection attendant upon classification and hypotheses selection. It appears that children and adults have a clear and a stable preference with respect to the speed with which they offer hypotheses in problem situations. Some children are fast; others slow.
>
> The reflectivity-impulsivity dimension describes the degree to which the child reflects upon the differential

> validity of alternative classifications of stimulus or solution hypotheses in situations in which many response possibilities are available simultaneously. In these situations some children have a fast conceptual tempo; they impulsively report the first classification that occurs to them or the first solution sequence that appears appropriate. The reflective children, on the other hand, characteristically delay before reporting a classification or solution hypothesis. They actively consider the alternatives available to them and evaluate their differential validity. The reflective child behaves as if he cared that his first response be as close to correct as possible."

The original work in this area by Kogan and his associates utilized three assessment techniques: Delayed Recall of Designs, Matching Familiar Figures, and the Haptic Visual Matching. Due to a crowded testing schedule it was decided to use two of these, omitting the Delayed Recall of Design.

The Haptic Visual Matching Test described earlier as a discrimination task was also timed with the purpose of using the time to respond as a measure of impulsivity-reflectivity.

In the Matching Familiar Figures Test the subject was shown a picture (the standard) and a four part choice array, only one of which was identical to the standard. The subject was asked to select the one figure from the array that was identical to the standard.

As with the Haptic the nature of the items were modified to make them less difficult. This was done by using differences that were more gross or discrete and reducing the choice array from six to four. The original procedure had used familiar objects as stimuli. The modified version used some geometric forms for early and easier items and then used some of the items of the original test, using familiar objects to complete the task.

The materials were assembled in loose-leaf form with corresponding standards and choice arrays on facing pages. As with the Haptic two measures were taken, the number of correct responses and the time from item presentation to response.

18. Illinois Index of Self-Derrogation

The Illinois Index of Self-Derrogation developed by Joseph H. Meyerowitz as a part of a larger project of the University Institute for Research on Exceptional Children, was the technique selected to elicit verbal self-estimates from children.

> "Due to the paucity of empirical research upon the self-concept of young children, normal as well as retarded, it was deemed necessary to develop an instrument. In developing an instrument, the working definition adopted was 'anything said about or attributed to one's self is, at least in part, one's self-concept'. This definition not only describes a significant manifestation of self-concept, but also permits objective measurement. The existence of an unexpressed, unmeasurable matric of components of the self-concept cannot be demonstrated, although few would deny its existence. However, there are responses which are put forth solely to influence favorably the behavior of others, because this is socially desirable; therefore these responses cannot be used to evaluate the subject's self-concept. These must be distinguished before this definition can be useful."

The original Illinois Index of Self-Derrogation was designed for administration to groups of five to seven children. However, the modifications in procedures utilized in the present study necessitated individual administration.

The IISD consists of 30 items, each item consisting of two sentences. One sentence describes a child with a socially undesirable ascription, such as, "Some children do not like the child with the balloon." The second sentence describes the child with a neutral or socially desirable ascription, e.g., "Many children like the child with the flag." In the original procedure the child listened to the descriptions of the two children and then marked in a designated area of his score sheet the child which he regarded as most like himself. In the present study this procedure was modified to accomplish two purposes: (a) remove the necessity of the child making marks and the dependency on verbal instructions and accuracy in following them, and (b) increase the child's manipulative activity in the task and hopefully increase the subject's task involvement and the subsequent validity of the responses. The score sheets of the original procedure had utilized stick figures (child with a balloon and child with a flag) that had been empirically shown to be equally attractive to children. These same stick figures were retained but were drawn on cards of heavy poster-board. The cards were approximately 4 x 4 inches and were yellow with the stick figures drawn in black. The two cards were placed in front of the child along with a brightly decorated box with an opening large enough to easily accomodate the insertion of the item cards. The child was then read a description of each child, e.g., "Some children do not like the child with the balloon. Many children like the child with the flag," and the child selected the card of the stick figure child which he regarded as most like himself and inserted it into the box. The examiner removed the remaining card and placed another pair of cards in front of the subject.

The ascribing of positive and negative characteristics of the balloon figure and the flag figure were systematically varied as were the placement (left-right) of the figures in front of the child, and the order (first, second) of the positive and negative statements. Of the 30 items on the scale two were warm up items, the remaining were treated as scorable derrogations. Six of these items were repeated as reliability check items.

As the numbers of positive and negative responses are complementary, either score format can be used. In this study results have been presented in terms of positive ascriptions rather than negative.

19. Delay of Gratification

An often stated characteristic used to differentiate the socio-economically disadvantaged is their propensity towards immediate gratification or inability to delay gratification. This assertion has been challenged by many scholarly sources, but valid or not, it remains as a part of the common perception of the lower economic class. The desirability of measuring delay of gratification was recognized since it has strong educational implications.

The technique selected was based on a procedure utilized by Walter Mischel (1961) in a study to investigate the effect of father-absence on children's ability to delay gratification. The relevant measure was taken to be preference for immediate, smaller reinforcement over delayed, larger reinforcement. According to the author:

> "Such choice preferences are fairly easily elicited and provide lifelike behavior which is readily quantified and some of whose empirical correlates have been explored independently. Namely, preference for delayed reinforcement (as opposed to immediate reinforcement has been related positively to accuracy in

> time statements and to social responsibility. . . .Preference for delayed reinforcement has also been negatively related to delinquency, positively to strength of "n Achievement", and negatively to acquiescence or "yeasaying" (as opposed to "nay-saying") tendencies. The preference for delayed reinforcement-immediate reinforcement distriction may thus be thought of as delineating two empirically-elaborated clusters, associated with significant differences in 'maturity', responsibility, delay over time, long term goal direction, and autonomy." (Mischel, 1961)

Mischel's various studies have all been conducted with children in the West Indies area of the Caribbean.

Following the procedure described by Mischel, 1¢ and 5¢ Tootsie Roll candy was chosen for the reward or reinforcement. At the completion of the testing of a child for the phase in question, the examiner would give the child a choice between a 1¢ Tootsie Roll immediately or a 5¢ Tootsie Roll to be received the next day the examiner returned to the class. The assessment consisted of recording the child's choice.

20. Family Relations Test

The Family Relations Test is an objective technique for exploring emotional attitudes in children. It was developed by Eva Bene and James Anthony, both of the Department of Psychiatry, the Institute of Psychiatry, London, England. The test is distributed by the National Foundation for Educational Research in England and Wales.

The technique was developed to fill clinical and research needs for objectively, reliably and rapidly indicating the direction and intensity of a child's feelings towards the various members of his family and of his (the child's) estimate of their reciprocal regard for him.

Two forms of the test are contained in the test materials. The authors do not specify an age range for either form, but only suggest that somewhere between ages 6 to 8, the child's level of comprehension should be adequate to employ the older form. The basic difference in terms of end results is that the younger form explores and classifies in terms of a broader attitude area than the older form. Given the ages of the sample in this study the younger form was used exclusively. The five classifications of responses that result are:

1. Positive feelings coming from child
2. Negative feelings coming from child
3. Positive feelings going towards child
4. Negative feelings going towards child
5. Dependence

The test materials are designed to concretely represent the child's family. It consists of 20 figures representing people of various ages, shapes and sizes, sufficiently stereotyped to stand for members of any child's family, yet ambiguous enough to become, under suggestion, a specific family. They range from grandparents to a perambulated baby, and from these the child is able to create his own family circle.

Another important figure incorporated into the test stands for "Nobody" and serves to accomodate those items that are not felt to apply to anyone in the family. Each figure is attached to a box-like base which has a slit in the top. The items are printed on small individual cards. On the back of the card is a symbolic representation of an address and stamp to give the item card an appearance of a post card which is to be mailed by placing the card in the slot of the box base for each figure. The child is told that the cards contain messages, and that their task is to mail each card "to the person" whom the message fits best.

The test situation is therefore basically a play situation and the test material is designed to prepare the subject for the emotional demands that are to follow.

The child's task is made easier by the test mechanics. He is not asked to verbalize analytically his many complex feelings for his family "on the spot" (the family constellation selected by the child in the test situation). This would lie outside the abilities and/or inclinations of most children. What he is expected to do, however, is to commit himself to a choice of pre-selected emotional attitudes, extracted from many different sources, but general enough to form a frame of reference for any child. The item therefore is "fixed" but its placement is free, and discarding into "Nobody" is permissible. Moreover, the feeling thrust into the figure immediately vanishes from sight leaving no incriminating trace. There is no visible reminder to the child of his distribution of love and hate, and, consequently, there is less guilt to interfere with his freedom of expression.

Scoring consisted of tallying the frequency with which items falling into the five categories were sent or attributed to the mother, father, siblings, self or Nobody.

APPENDIX B

This appendix contains item difficulty data in the form of mean score on all items in the Wechsler Pre-School and Primary Scale of Intelligence for the five groups of children studied. Sample sizes are given in the Information subtest section.

Information subtest (scored 1 or 0)

		Group (N in parentheses)			
Item	A(34)	B(20)	C(26)	D(26)	E(29)
1. Nose	1.00	1.00	1.00	1.00	1.00
2. Ears	1.00	.95	.96	1.00	1.00
3. Thumb	.85	.80	1.00	1.00	1.00
4. Bottle	.85	.90	.92	.88	.93
5. Live-water	.82	.50	.96	1.00	.93
6. Grass	.82	.90	.96	.96	1.00
7. Animals (3)	.74	.70	.77	.92	.97
8. Milk	.91	.70	1.00	1.00	1.00
9. Shines-night	.44	.40	.77	.81	.79
10. Legs-Dog	.58	.45	.88	.92	1.00
11. Letter-Mail	.38	.40	.65	.69	.90
12. Wood	.69	.40	.88	.69	.79
13. Round (2)	.44	.40	.73	.77	.86
14. Water-boil	.53	.40	.65	.88	.83
15. Store-sugar	.56	.45	.62	.58	.45
16. Pennies	.65	.30	.23	.31	.31
17. Shoes	.09	.10	.04	.31	.45
18. Days-week	.03	.05	.12	.15	.41
19. Bread	.12	-	.15	.31	.34
20. Seasons	-	-	.12	.19	.41
21. Rubies	.24	.10	.12	.12	.69
22. Dozen	-	-	-	.08	.69
23. Sun-set	.06	-	.04	-	-

Vocabulary subtest (scored 2, 1 or 0)

			Group		
Item	A	B	C	D	E
1. Shoe	1.18	1.15	1.50	1.00	1.62
2. Knife	1.62	1.40	1.84	1.81	2.00
3. Bicycle	1.06	.85	1.08	.88	1.07
4. Hat	1.41	1.45	1.58	1.42	1.66
5. Umbrella	1.32	1.10	1.54	1.50	1.59
6. Nail	.88	.55	1.04	1.23	1.34
7. Letter	1.23	1.00	1.27	1.57	1.62
8. Gasoline	.79	.55	1.04	.88	1.21
9. Donkey	.44	.35	.85	.65	.72
10. Swing	.38	.25	.42	.58	.45
11. Castle	.41	.20	.81	.65	.80
12. Snap	.47	.60	.85	.73	.93
13. Fur	.12	.25	.50	.65	.72
14. Polite	.09	.05	.92	.69	1.24
15. Moth	.03	.15	.35	.42	.89
16. Join	.09	.15	.35	.31	.55
17. Hero	.06	.05	.35	.38	.62
18. Diamond	-	.15	.23	.50	.55
19. Chisel	-	.05	.08	.15	.21
20. Nuisance	-	-	.15	.04	.69
21. Microscope	-	-	-	.04	.17
22. Gamble	.09	.05	.12	.04	.07

Arithmetic subtest (scored 1 or 0)

Item	Group				
	A	B	C	D	E
1.	1.00	.90	1.00	1.00	1.00
2.	1.00	.80	1.00	1.00	1.00
3.	.97	.80	1.00	.96	1.00
4.	.65	.65	.88	.96	.93
5.	1.00	.80	.96	.96	1.00
6.	.88	.70	1.00	.92	1.00
7.	.79	.45	.88	.92	1.00
8.	.85	.55	.92	.92	.93
9.	.71	.65	.73	.81	.93
10.	.41	.30	.73	.88	.90
11.	.41	.35	.58	.81	.79
12.	.26	.20	.38	.54	.52
13.	.18	.20	.42	.61	.72
14.	.15	.10	.19	.42	.52
15.	.12	.05	.23	.23	.28
16.	.29	-	.12	.19	.17
17.	.12	-	.15	.27	.31
18.	-	.05	.04	.04	.07
19.	-	.05	-	.04	.10
20.	-	-	-	.12	-

Similarities subtest (scored 1 or 0)

Item		A	B	Group C	D	E
1.	Train	.85	.55	.85	.88	.76
2.	Shoes	.97	.85	1.00	.96	.97
3.	Ball	.62	.65	.73	.77	.83
4.	Glass	.74	.55	.77	.88	.86
5.	Bread-Meat	.58	.55	.69	.65	.83
6.	Legs	.65	.60	.73	.85	.76
7.	Pencil	.47	.45	.69	.73	.73
8.	Boys-men	.65	.50	.81	.88	.79
9.	Milk-water	.41	.45	.69	.69	.79
10.	Knife-glass	.09	.25	.19	.23	.45
11.	Coat-sweater	.29	.25	.38	.50	.48
12.	Piano-violin	.26	.30	.77	.73	.93
13.	Plum-peach	.47	.30	.50	.58	.76
14.	Penny-nickel	.21	.35	.62	.58	.97
15.	Beer-wine	.12	.25	.19	.38	.55
16.	Cat-mouse	-	.05	.15	.15	.34

Comprehension subtest (scored 0, 1 or 2)

Item		A	B	Group C	D	E
1.	Play-matches	1.38	.95	1.35	1.62	1.72
2.	Wash	.94	1.05	1.38	1.35	1.41
3.	Cut-finger	1.47	1.06	1.85	1.92	1.69
4.	Clocks	1.47	1.15	1.65	1.69	1.86
5.	Baseball-(doll)	.97	.65	.88	.96	.62
6.	Toilet	1.21	.85	1.46	1.58	1.58
7.	House-windows	.74	.65	1.04	.88	1.21
8.	Clothes	.38	.20	.65	.54	.72
9.	Work	.56	.80	1.04	.81	.59
10.	Light-room	.35	.30	.50	.58	.76
11.	Children-sick	.15	.15	.58	.69	1.06
12.	Loaf-bread	.76	.40	1.07	1.12	.93
13.	Light	.03	.25	.58	.46	.79
14.	House-brick	.06	.10	1.92	.35	.62
15.	Criminals	.12	.25	.38	.19	.41

Picture Completion subtest (scored 0 or 1)

			Group		
Item	A	B	C	D	E
1. Comb	.88	.85	.92	.96	1.00
2. Wagon	.94	.80	.81	.88	1.00
3. Doll	1.00	.85	.96	.88	.97
4. Roses	.94	.75	.92	.85	.97
5. Girl	.74	.45	.92	.65	.93
6. Fox	.74	.65	.85	.92	.90
7. Table	.82	.70	.96	.92	1.00
8. Seesaw	.50	.45	.54	.85	.83
9. Hand	.50	.45	.81	.77	.83
10. Cat	.53	.50	.65	.77	.90
11. Bridge	.38	.40	.62	.62	.79
12. Clothesline	.56	.50	.50	.62	.79
13. Watch	.32	.30	.38	.54	.52
14. Shoes	.44	.30	.50	.50	.66
15. Automobile	.21	.15	.42	.54	.69
16. Swing	.29	.15	.35	.58	.72
17. Door	.24	-	.38	.35	.55
18. House	.29	.20	.35	.46	.48
19. Coat	.12	-	.15	.23	.24
20. Card	.21	.15	.19	.19	.24
21. Rooster	.15	.05	.19	.08	.21
22. Scissors	.09	.10	.04	.12	.17
23. Screw	.06	-	-	.04	.14

Mazes subtest (scored 0, 1 or 2)

			Group		
Item	A	B	C	D	E
1. (0-2)	1.88	1.80	2.00	1.97	1.97
2. (0-2)	1.65	1.50	1.92	1.92	1.90
3. (0-2)	1.00	1.00	1.58	1.58	1.69
4. (0-2)	.74	.95	1.62	1.65	1.48
5. (0-2)	.56	.65	1.42	1.58	1.59
6. (0-2)	.62	.50	1.42	1.42	1.59
7. (0-3)	.24	.45	1.50	1.46	2.07
8. (0-3)	.06	.25	1.08	1.31	1.76
9. (0-4)	.09	.25	1.15	.85	1.38
10. (0-4)	.06	-	.65	.46	1.31

Geometric Design subtest (scored as indicated under Item)

Item		Group A	B	C	D	E
1.	(0-2)	1.59	1.50	1.88	1.73	1.66
2.	(0-2)	1.24	1.05	1.23	1.31	1.52
3.	(0-2)	1.35	1.15	1.27	1.19	1.45
4.	(0-2)	.85	.64	1.15	1.12	1.48
5.	(0-2)	1.09	.90	1.35	1.23	1.59
6.	(0-3)	.97	.75	1.08	1.50	1.69
7.	(0-3)	.94	.50	1.04	1.35	1.76
8.	(0-4)	.68	.60	.54	1.38	2.21
9.	(0-4)	.24	.15	.46	.65	1.52
10.	(0-4)	.18	.30	.27	.96	1.38

Block Design subtest (scored 0, 1 or 2)

Item	Group A	B	C	D	E
1.	2.00	1.85	1.92	1.92	1.86
2.	1.88	1.80	2.00	1.96	1.83
3.	1.29	1.30	1.50	1.77	1.55
4.	1.15	1.15	1.58	1.77	1.48
5.	1.18	1.15	1.50	1.46	1.52
6.	.44	.70	.65	1.35	1.24
7.	.74	.80	1.00	1.42	1.52
8.	.68	.65	.77	1.35	1.45
9.	.06	.05	.15	.15	.59
10.	.15	.35	.19	.35	.72